

*Marcelo de Araujo**

Elucidating the narrative structure of arguments for moral contractarianism

Abstract

If the nature of an agreement is such that it is binding upon the agreeing parties, then before an agreement is made the parties to the agreement have to share the common understanding that, by agreeing, they have an obligation to act as agreed. But where does that shared understanding come from? Does the shared understanding simply dawn on them the moment they agree to do something, or do they already share that understanding before they agree? The assumption that the very idea of an agreement cannot be established by means of an agreement and that, for this reason, social contract theories have to rely on at least some basic normative ideas for the establishment of less basic normative ideas has been a common line of criticism against moral contractarianism. It does not follow from this, though, that supporters of contractarianism cannot answer this objection. In this paper, I argue that contractarianism can, indeed, tackle this problem.

Keywords

agreements; contractarianism; narratives; Adolf Reinach; social acts

Introduction

If the nature of an agreement is such that it is binding upon the agreeing parties, then before an agreement is made the parties to the agreement have to share the common understanding that, by agreeing, they have an obligation to act as agreed. But where does that shared understanding come from? Does the shared understanding simply dawn on them the moment they agree to do something, or do they already share that understanding before they agree? The assumption that the very idea of an agreement cannot be established by means of an agreement and that, for this reason, social contract theories have to rely on at least some basic normative ideas for the establishment of less basic normative ideas

* Federal University of Rio de Janeiro and State University of Rio de Janeiro.

has been a common line of criticism against moral contractarianism.¹ It does not follow from this, though, that supporters of contractarianism cannot answer this objection. In this paper, I argue that contractarianism can, indeed, tackle this problem.

I firstly advance three fictional situations depicting the gradual emergence of agreements among rational agents. Then I argue that the narrative structure underlying these fictional situations are not simple illustrations for the purpose of an inquiry into the concept of agreement. They constitute, rather, the method that guides the inquiry itself. We literally imagine a narrative in which the concept under scrutiny – the concept of agreement – did not yet exist and ask if, starting from there, we would have reasons to arrive where we are. In the last section of this paper, I show, then, that although the three fictional narratives I propose do not purport to be a historical account on the emergence of agreements, they are in line with a method of investigation that is not strange to other fields of inquiry. I argue, indeed, that the narrative structure of arguments for moral contractarianism must be at least compatible with all we know about our evolutionary history, or the history of human languages, and human institutions at large.

1. Before the Agreement

What are people doing when they say that they *agree* about something? Sometimes, they simply aim at avoiding the repetition of a sentence that has just been uttered by someone else. In these cases, *agree*, as a verb, plays the same role the adjective *true* often plays when applied to a sentence. If A makes a statement about something and B says he agrees with A, B gives to understand that he acknowledges that a sentence has been

¹ See Mapel, “The contractarian tradition and international ethics,” 183: “As critics outside the tradition have often pointed out, contractarians must rely on extra-contractarian assumptions in order to explain the moral force of the contract. For example, Locke’s social contract presupposes that individuals already possess certain natural rights. Those rights are not justified in contractarian terms; rather, they explain the moral force and scope of the contract itself.” See also Larmore, *The Morals of Modernity*, 130; Araujo, “Kontraktualismus”. In this paper, I am not concerned with *contractarianism* as a political theory, but as a moral theory. Neither am I concerned in this paper with *contractualism*, that is a kind of social contract theory that, unlike *moral contractarianism*, does not try to explain the normative force of basic concepts such as, for instance, “agreement”, “contract”, or “right” solely on contractual grounds. John Locke, John Rawls, and Thomas Scanlon, to name just three authors, can be seen as philosophers in the tradition of *contractualism*. David Gauthier and Peter Stemmer, on the other hand, can be seen as supporters of *moral contractarianism*.

uttered by A and that he, like A, accepts *that* sentence as true.² Sometimes, though, by agreeing to what one person A says, another person, B, commits himself to the enactment of an action, namely the action that has been proposed, or suggested, or recommended by person A. But what exactly is at stake when B *commits* himself to what A proposes, or suggests, or recommends remains to be explained. Consider, for instance, these three hypothetical situations depicting a sequence of interactions between persons A and B.

1.1. SITUATION 1

It is a sunny day. A and B go out on a hunt. A utters the proposition p “It’s a lovely day today”, to which B replies by simply repeating p . In saying that p , B shows that he, too, believes that p . But B might as well have reacted with one of these less repetitive utterances:

1. “Yes.”
2. “Indeed.”
3. “It is true.”
4. “I agree with you.”

In saying that he *agrees* with A, B does not commit himself to anything other than the recognition that A has uttered p and that he, like A, also believes that p is true. If B changes his mind ten minutes later and states $\sim p$, because he realizes now that he would actually prefer it were raining, he will give A to understand that he believes p is false. In this case, A would not be entitled to blame B for having changed his mind, that is for not sticking to what he had said ten minutes earlier, namely that p . B agreed with A about what was the case in the world, or how he felt about the world, as it seemed to him then, but upon reflection B may change his mind and come to believe that p is false, even while also admitting that p may remain true from A’s perspective. B may of course have those beliefs about the weather while failing to express them. If B remains silent, and yet unconsciously and unintentionally behaves in such a way that A can reasonably infer that B thinks it is a lovely day, A can reasonably assume – all thing being equal – that B is in *agreement* with A about the weather. We will see later, though, that an *agreement*, in the relevant sense of being binding on both A and B, cannot be inferred in a similar way.

It should be noted that I assume here, for the sake of the argument, that B did not utter p with the sole intent of being friendly towards A. In saying that p , and then ten minutes later $\sim p$, B is truthfully expressing his beliefs about the world, or how he feels about the world.

² Kirkham, *Theories of Truth*, 320-21.

Sometimes, though, maybe indeed on most occasions, when we utter a proposition like p , we do not really intend to express our opinion about the world, or our real feelings about the weather; we generally aim at things more prosaic such as, for instance, being friendly towards the person we are talking to regardless of our actual opinions about how things are in world.³ The propositions uttered in the brief exchange between A and B, as depicted in situation 1, might be easily taken as an example of such non-declarative sentences. But as I have already explained, I assume here that, in saying that he agrees with A, B does not intend to be friendly towards A. B wants A to know that he too, like A, thinks it is a pleasant day today.

A and B have a long day ahead of them. The sky is gradually turning grey. It seems it is going to rain.

1.2. SITUATION 2

Two hours have elapsed since A and B have last talked about the weather. The rain is falling, and it is going to be dark soon. Neither B nor A want to get a cold while spending the night in the woods. A utters the proposition q to B:

“Let us build a shelter together.”

Unlike p , q is not a declarative sentence. It is an imperative sentence, which A utters to make a proposal, or a suggestion, or a recommendation, or an exhortation to B. The nature of the relationship between A and B is such that neither of them is in a position to issue an order or a command to the other. And the nature of the task proposed by A is such that neither A nor B can build a shelter alone. They have to cooperate for the purpose of building a shelter together and, thus, avoid the nuisance of becoming wet and, possibly, ill. B replies to A's proposal with r :

“*I like* this proposal. q .”

It is important to note now that although the sentence q can be neither true nor false (whether uttered by A or B), the sentence “I like this proposal” is a declarative sentence. It is a statement about what is going on in B's mind, and it is true to the extent that B really appreciates the prospect of building a shelter with A. Thus, the sentence “I like this proposal” is not true because B decided to tell A that he likes the idea of building a shelter. It is rather the other way round: it is because the sentence is true that B decides to tell A this. If B were speaker of a romance language such as Portuguese, instead of saying that he *likes* A's proposal, he might as

³ One study from 2006, for instance, suggests that on average only 36 percent of propositions uttered in the course of a six-minute informal conversation are declarative sentences. See Jurafsky, “Pragmatics and computational linguistics.”

well have said: “This proposal *agrada-me. q*”. But let us suppose, alternatively, that B is a speaker of Middle English. Back in 1350, B might have reacted to A’s proposal by saying “This proposal *agree me. q*.” Three different sentences – one in modern English (“I like this proposal. *q*”), one in Portuguese (“This proposal *agrada-me. q*”), and one in Middle English (“This proposal *agree me. q*.”) – have the same function: they enable B to let A know that B appreciates A’s proposal, namely to build a shelter together. The first verb in the second and third sentences (*agrada* and *agree*) are etymologically related to each other. In using the verb *agree* to express his appreciation for A’s proposal in Middle English, B is not yet committing himself to join A in the cooperative effort to build a shelter. B is simply saying that A’s proposal is *agreeable* to him. The *Oxford English Dictionary* cites this line from Geoffrey Chaucer’s *Troilus & Criseyde* (ca. 1380) as the earliest known occurrence of the verb *agree*: “If harme agree me wher-to pleyne I þanne.”⁴ This has been rendered in Modern English as “If harm is agreeable to me, why do I lament?”⁵ Because B explicitly said that A’s proposal is agreeable, A is right in assuming that B will help A throughout the process of addressing a common problem, namely lack of protection against the rain during the night. But does it also mean that A has a legitimate claim, or has the right, or is entitled to require B to stick to what he has said, even if the task proves less agreeable to B later? What if B (again) changes his mind and then, ten minutes later, he realizes that the work does not please him, that building a shelter with A is not an agreeable task after all, and that he would rather endure the rain in his thick leather coat rather than having the trouble of building a shelter in the dark?

One might suggest that B is under an obligation to help A because in saying that he *liked* the prospect of building a shelter with A, B implicitly committed himself to the task of cooperating with A. According to this understanding of what is at stake in Situation 2, when B says that a joint undertaking is *agreeable* to him, he implicitly means something else, namely that he will stick to what he says by behaving in the future *as though* the prospect of cooperating with A will remain no less attractive or agreeable to him than it seemed at the beginning, when B first uttered *r*. According to this understanding of Situation 1, one might further suggest that there was an *implicit agreement* between A and B, binding on both of them, so that A is entitled to count on B’s help even if B later realizes that the task is less agreeable to him than he had initially thought.

⁴ See “agree, *v.*” In *OED (Oxford English Dictionary) Online*. June 2021. Oxford University Press. See also Skeat, “Agree (*v.*)”, 11; Klein, “Agree (*Intr. v.*)”, 40.

⁵ Chaucer, *Troilus and Criseyde*, 10.

However, it seems to me that this is a wrong understanding of what is at stake in Situation 2.

When A hears *r* from B, A has *prima facie* good reasons to assume that B will cooperate with him throughout the entire process of building a shelter. B makes a statement about what is going on in his own mind, and A (reasonably) accepts B's statement as true (call it *assumption 1*). But A has to rely on a further assumption for believing that B will cooperate throughout the entire process of building a shelter, namely that B will not change his mind about the attractiveness of their joint undertaking in the same way B changed his mind about the weather in Situation 1 (call it *assumption 2*). This is a reasonable assumption as well, for both A and B have an interest in not getting wet and, possibly, ill. As it happens, though, people do change their minds sometimes, especially due to incomplete information about projects they venture to undertake, whether on their own or in partnership with other people. If B later changes his mind about the attractiveness of the joint undertaking with A and, accordingly, decides not to cooperate with A any longer, all that A can do is to express his own disappointment towards B. Depending on how disappointed A is, A can make a step further and vent all his frustration and anger at B. But A's anger, disappointment, and frustration by themselves do not entitle him to require B to stick to *r*. The feelings A has towards B do not endow A with any special rights vis-à-vis B. Person A could only be entitled to require B to stick to *r*, if A had good reasons to rely on yet another assumption, namely that B will cooperate throughout the entire process of building a shelter whether or not this task becomes less agreeable to B later (call it *assumption 3*). But why would A make *assumption 3*? What if A and B had never engaged in the practice of behaving in the present *as though* their behaviour were driven by a motivation that they may not have anymore – a motivation that can be expressed in modern Portuguese and Middle English with sentences containing cognate verbs such as *agradar* and *agree* respectively? If A and B have never engaged in the practice of acting *as though* they had a motivation that, as a matter of fact, they may not have anymore, they have good reasons to give it a try and see what happens.

It would be mutually advantageous for A and B, then, to engage in a practice like this: when B utters *r*, A is entitled to understand that B will behave accordingly in the future, even if the motivation that B had to utter *r* in the first place becomes weaker, or disappears later. Before the emergence of this new practice, A and B used the verb *agree* primarily to express what was *agreeable* to them. In the new practice, the verb retains its original meaning: B agrees to build a shelter with A because the prospect of doing that is agreeable or attractive to him, otherwise he would not agree in the first place. But in the new practice the verb would

acquire a nuance it did not have before. In saying that he *agrees* to build a shelter with A, B would not be expressing what he thinks is agreeable to him now. The sentence containing the verb *agree* would not be a statement about what is going on in B's mind. In saying that he *agrees* to build a shelter with A, B would be committing himself to the performance of an action. This means that B would give to understand that if the performance of that action fails to be agreeable to B later, B would perform it anyway, *as though* he felt that the prospect of building a shelter with A were no less attractive than it seemed before.

The existence of the new practice would be advantageous for A, for A needs B to build a shelter. But it would be equally advantageous for B, for B, too, is likely to feel disappointed, frustrated, and angry at A, if A changes his mind about the attractiveness of joint projects that B has already started to undertake under the assumption that A will cooperate with him. However, A's and B's implicit reasons to create and maintain the new practice, which would regulate their behaviour towards each other in the future, do not automatically give rise to implicit agreements in the minds of either A or B. Both A and B may have the implicit intention of co-creating and engaging in the new practice and acting accordingly. A (and it also holds for B) may even have some good reasons to believe that B has already figured out that it would be mutually advantageous for them to create the new practice. Yet, A cannot infer from this that there is an *implicit agreement* between him and B. The implicit intention of creating a new practice is not yet that practice and cannot produce the same effects of that practice. If B utters *r*, but does not act accordingly later, A cannot simply say that he knew that B knew perfectly well that it would be mutually advantageous for them if there were a binding agreement between them and that, for that reason, A is entitled to require B to act as *agreed*. B did not agree *about* anything other than this: that it would be mutually advantageous if the new practice existed. But from this A cannot infer that B implicitly agreed *to* the enactment of an action, namely: to build a shelter with A. Biding agreements must exist in the world before they can exist as implicit agreements in A's and B's minds.

B can privately *agree* with A about how things are in the world, or how he feels about the world, and yet refrain from stating this because, for instance, it would be too embarrassing for B to admit that he actually *agrees* with A. Maybe A has already figured out that B is actually in *agreement* with him, even though B will not state that explicitly. But B cannot have an *agreement* – a binding agreement – with A without addressing A and saying the relevant words – or making the relevant gestures – by means of which B *entitles* or *empowers* A to require something from B, in case B does not act *as agreed*. Implicit agreements can eventually come into

existence, but they are a further development of a practice that has yet to emerge between A and B. For now, though, this practice does not yet exist between A and B.

In order to see this point more clearly we can alternatively think of A and B as two powerful states eager to avoid a nuclear war between them. Both President A and President B have already figured out that it would be mutually advantageous not to strike first because one cannot strike and still reasonably expect that the other will not retaliate. President A knows that, and he knows that the President B knows that as well and, for this reason, President A thinks to himself: “Deep inside President B *agrees* with me that it is mutually advantageous not to strike first”. Then President A elaborates a bit further and thinks to himself: “Deep inside President B is in *agreement* with me as regards the truth of this sentence: ‘it is mutually advantageous not strike first’ for both of us will perish in the event of a nuclear war” (call this *agreement*₁). Yet – all things being equal – there is no implicit *agreement* between A and B as regards committing themselves to the action of refraining from striking first (call this *agreement*₂). *Agreement*₁ is an epistemic agreement: it is an agreement *about* how things are in the world, or how things are in other people’s minds. *Agreement*₂, on the other hand, is a normative one: it is an agreement *to* the pursuit of peace rather than war.

Now, what if both President A and B *explicitly* manifest their intention to pursuit peace, rather than war? Has an *agreement*₂ arisen among them now? Appearances to the contrary, in this case, too, all things being equal, the sheer declaration of will does not yet constitute a commitment to the pursuit of peace rather than war. The declaration of will is primarily a statement about what one actually wills. It is not a commitment (or an *engagement*, as for instance H.A. Prichard uses the word).⁶ Just in the same way a promise is not a statement about what one intends to do in the future, but a unilateral commitment of the promisor towards the promisee, an agreement, in like manner, is not a statement about how the agreeing parties intend to act, but a commitment to some specific kind of behaviour towards each other.

One philosopher who was particularly clear about this point is Adolf Reinach. As early as 1913, Adolf Reinach argued that acts such as commands, requests, promises and so on are not a “declaration of will” (*Willensäußerung*). These are “social acts” – or “speech acts”, as John Austin and John Searle would call them some fifty years later. *Agreement*₁ would not count as a “social act” because this kind of agreement can exist without one party B having to address the other party A in such a way that

⁶ Prichard, ‘Does Moral Philosophy Rest on a Mistake?’, 3.

A takes notice of what B is doing, namely *being in agreement*₁ with A. Indeed, we can imagine, as I have suggested earlier, that B may actually want to hide from A that B is in agreement with A. *Agreement*₂, on the other hand, is a “social act” because it can only exist as a commitment if one party B addresses the other party A and makes sure that A understands the relevant words or, as the case may be, the relevant gestures by means of which things such as commands, requests promises etc are brought into existence. According to Reinach, “social acts” are, in this regard, necessarily “in need of being heard” (*vernehmungsbedürftig*).⁷ If B does not succeed in making A to take notice of the relevant words, or the relevant gestures, B’s attempt to produce the relevant act “misfires”, that is B fails to bring about the relevant state of affairs, whether it is a command, a request, a promise etc. As Reinach aptly puts it: “It can of course happen that commands are given without being heard. Then they fail to fulfil their purpose. They are like thrown spears which fall to the ground without hitting their target.”⁸ It is *in the performance itself* (*im Vollzug selbst*) of the act of addressing one party, and of making oneself noticed and understood by the other party, that one party produces the relevant social effect.⁹ Now, it might be argued, then, that if A and B explicitly let each other know what is in their minds concerning their respective wills, they are producing a social act, namely a kind of *agreement*₂. It seems to me though, that Reinach correctly realized that by explicitly revealing to one another what is in their respective minds, A and B are not yet agreeing to the performance of an action, they are rather *informing* each other what they intend to do. Although *informing* is, indeed, a “social act” it differs in some important respects from the social act of promising and,

⁷ See Reinach, *The a priori Foundations of the Civil Law*, 19; Reinach, *Die apriorischen Grundlagen des bürgerlichen Rechts*, 707. The English verb *hear*, in John Crosby’s translation, is too restrictive, as the corresponding German verb *vernehmen* can also be used to denote *taking notice* of something by seeing it, or touching, or smelling etc.

⁸ See Reinach, *The a priori Foundations of the Civil Law*, 19; Reinach, *Die apriorischen Grundlagen des bürgerlichen Rechts*, 707. The corresponding German text reads: “Der Befehl ist seinem Wesen nach *vernehmungsbedürftig*. Wohl kommt es vor, daß Befehle erteilt, aber nicht vernommen werden. Dann haben sie ihre Aufgabe verfehlt. Sie sind wie geschleuderte Speere, welche niederfallen, ohne ihr Ziel zu erreichen.” (Italicized in the original).

⁹ See Reinach, *The a priori Foundations of the Civil Law*, 19-20: “What has been shown for commanding also holds for requesting, warning, questioning, informing, answering and for still many other acts. They are all social acts, which, by the one who performs them and *in the performance itself*, are as it were cast towards another person in order to fasten themselves in his soul.” The corresponding German text (p. 707) reads: “Was hier für den Befehl ausgeführt wurde, gilt auch für das Bitten, Ermahnen, Fragen, Mitteilen, Antworten und noch vieles andere. Sie alle sind soziale Akte, welche von dem, der sie vollzieht, *im Vollzuge selbst* einem anderen zugeworfen werden, um sich in seine Seele einzuhacken.” (Italicized in the original).

for the purpose of my investigation in this paper, from acts such as *agreements*₂. Let me quote Reinach's point in full length:

We now see clearly how thoroughly mistaken and untenable is the usual conception of promising as an expressing of intention or of will. An expression of will runs like this: I intend. If it is directed to someone, then it is an informing, which is indeed a social act but no act of promising. And of course it does not become a promise by being directed to the one who will profit from the intended action. Promising is neither intending nor the expression of intending; it is rather an independent spontaneous act which in turning without, expresses itself.¹⁰

It seems to me that Reinach's insightful analysis of social acts such as commands, requests, promises etc. is valid for *agreement*₂ as well. If A figures out that deep inside B is in *agreement*₁ with A as to the advantages of *agreements*₂, A is not thereby entitled to assume that B has agreed to do anything. But neither has B agreed to do anything if he explicitly says "I intend" (*Ich will*) to pursue this course of action because that is agreeable to me. There emerges the question, then, as to what makes the transition from a simple declaration of will to a *agreement*₂ possible. Something is missing and at this juncture my own analysis departs from that proposed by Reinach. While Reinach intends to spell out the *a priori* nature of normative concepts such as promises and contracts, regardless of the institutional framework in which they actually exist, I argue, instead, that it is the institutional framework in which they exist that endows concepts such promises, contracts and – more importantly – agreements with the normative force they have. What is missing is A's capacity to counteract with something disagreeable to B, if the prospect of building a shelter with A becomes less agreeable to B later.

In the end, A got ill due to the bad weather and B returned empty-handed from the woods, for A was too weak to help B to set up the traps. They would never go hunting together again.

¹⁰ Reinach, 26. The original German text (p. 715) reads: "Wir sehen jetzt klar, wie gänzlich schief und unhaltbar die übliche Auffassung des Versprechens als einer Vorsatz- oder Willensäußerung ist. Eine Willensäußerung lautet: Ich will. Sie kann sich an jemanden wenden, dann ist sie eine Mitteilung, ein sozialer Akt zwar, aber kein Versprechen. Und auch dadurch wird sie natürlich nicht zum Versprechen, daß sie sich an denjenigen wendet, in dessen Interesse das vorgesezte Verhalten liegt. Das Versprechen ist weder Wille, noch Äußerung des Willens, sondern es ist ein selbständiger spontaner Akt, der nach außen sich wendend, in äußere Erscheinung tritt."

1.3. SITUATION 3

Three hundred years have elapsed since the early death of A and B. But their families have remained on good terms in spite of occasional rifts in the past. The weather, though, did not change much since Chaucer's days. A and B decide to go on a hunt, as some of their ancestors used to do. But then (again) it started to rain. They have now a wider range of tools to build a shelter, but the task still requires a lot of cooperative work. A (again) takes the lead and utters *q*: "Let us build a shelter." This time, though, instead of saying that A's proposal is *agreeable*, B replies with the sentence *s*:

"I *agree* with your proposal, my dear A. *q*."

But then, ten minutes later, B (again) changes his mind and realizes that building a shelter is too demanding a task. B wears a waterproof coat, much thicker than A's, and now realizes that getting a bit wet is a minor nuisance compared to the task of building a shelter in the dark. B does not care that A has already had the trouble to spot a suitable place and collected some logs for the shelter. Is A now entitled to blame B for not sticking to what he said earlier, namely *s*? In situations 1 and 2, B changed his mind but – apart from A's disappointment – the world remained the same. But now, in changing his mind, B also changed the world: B *broke* an agreement that existed between him and A.

In order to build shelters more effectively, successive generations of As and Bs devised new tools: sharper saws, stronger ropes, lighter nails and so on. But the tools by themselves did not enhance cooperation, which they gradually realized was equally important for the purpose of building a shelter – or maybe even more important than the other tools. Thus, the As and Bs gradually devised a cooperation-enhancing tool to enforce cooperation in the face of fading motivation: the institution of *agreements*, (henceforth simply agreements). Like other tools, agreements can be thought of as another kind of artefact that A and B have brought about *in the world*.¹¹ Now, in uttering *s*, B commits himself to cooperating with A. This means that B thus gives A to understand that he will act accordingly whether or not the motivation he has now becomes weaker or disappears later. And A understands this because, in uttering *s*, B is actually *empowering* A to exact the relevant behaviour from B, if B fails to act as agreed. There emerges the question, then, about what hap-

¹¹ See Reinach, 22: "If a command is given or a request is made, something is thereby changed in the world." The corresponding German text (p. 711) reads: "Ist ein Befehl oder eine Bitte vollzogen, so hat sich damit etwas geändert in der Welt." For an account of social norms as kinds of "artefacts", see e.g. Stemmer, *Handeln zugunsten anderer*, 201, 255; Stemmer, "Moralische Rechte als soziale Artefakte."

pens if the original motivation does indeed disappear in the future: what will prompt B to perform the action he gave to understand he would perform, if B does not feel like doing it anymore? This is a pressing question for A, for how can A be reasonably sure that B will stick to *s* in the face of fading motivation?

The cooperation-enhancing tool developed and tested by previous generations of As and Bs is activated when either A or B uses the verb *agree*, as it occurs in *s*. In order to activate the cooperation-enhancing tool, B has to *address* B and utter the relevant words: “I *agree* with your proposal, my dear A. *q*.” – one can of course drop the vocative. The cooperation-enhancing tool will not be activated if B just (privately) thinks the relevant words, or utters them in a language A cannot understand. B of course is not compelled to activate the cooperation-enhancing tool, if the prospect of cooperating with A does not appear agreeable to him. But there is a crucial difference now: If the prospect of cooperating with A seems agreeable to B now, and if for this reason B decides to activate the cooperation-enhancing tool, the prospect of *not* cooperating with A, after B has already activated the tool, will be very disagreeable to B later. B may fail to have the original motivation later, but he will still have a strong motivation to avoid some disagreeable things that may happen to him, if he fails to act as agreed. Both A and B are aware that the tool works according to this simple rule. The tool also admits of varying degrees of disagreeableness to those who activate it and do not act accordingly. Some of these quite disagreeable things may happen to B, if B does not act as agreed:

- Lack of reputation as a good co-operator before A and other people.
- Being subjected to ostracism by A and other people.
- Being excluded from future cooperative projects with A and other people.
- Being subjected to informal retaliation (e.g. other people will gossip about B).
- Being subjected to formal retaliation (e.g. B will have to pay a fine, or be hanged by the neck).

Of course, this is not an exhaustive list. And of course, as with other human tools, the cooperation-enhancing tool may sometimes fail to work as expected. B may still break the agreement and nonetheless, some way or other, manage to evade the disagreeable things he was supposed to experience if he fails to act as agreed. But as long as the cooperation-enhancing tool works more often than it fails, the As and Bs, in the course of generations, have more reasons to keep the tool than they have to throw it away.

The cooperation-enhancing tool shares with ordinary tools such as hammers, nails, and ropes another important feature. Some ordinary tools come into existence as the result of well-thought design and careful planning. They do not simply emerge – they are invented, created etc. Some tools, on the other hand, come into existence in a different way: we discover, more or less by chance, a useful function in an object that was not originally designed to fulfil that function – or that was not designed at all. A piece of bone, for instance, is not a weapon, but in casually figuring out that a piece of bone may work as a weapon, one may want to keep it for later use and see what happens.¹² By way of small incremental changes, and without much planning, weapons such as bludgeons, spears, and arrows will gradually *emerge*.¹³ Sophisticated tools can also be repurposed to fulfil functions they had not been designed for. The substances behind contraceptive pills and Viagra, for example, were originally conceived to address menstrual disturbance and heart disorder, respectively. But researchers discovered, as they tentatively deployed the drugs, that these substances caused side effects that were quite desirable, even though the researchers did not aim at the side effects when they first designed the drugs. Thus, they kept the drugs and repurposed them to fulfil new functions.

Now, it is clear that that institutional tools, too, can be created as the result of well-thought design and careful planning, but they can also simply *emerge*. In this case, the rules that underlie the use of the tool may be an afterthought to the very practice of *using* the tool. The rules may be figured out on a tentative basis, so to speak, “as we go along”, in the same way the rules of a simple game may be established as the game gradually evolves.¹⁴ The rules that underlie the use of the cooperation-enhancing tool are quite simple: B activates the tool by addressing A and uttering that he *agrees* to perform an action with A. B is aware that if the prospect of cooperating with A looks agreeable to B now, the prospect of *not* cooperating with A will be disagreeable to B later. Once B has activated the tool, B cannot deactivate it on his own. It takes both A and B to deacti-

¹² I have in mind here a scene from 1968 Stanley Kubrick’s *2001: A Space Odyssey*. The scene is available at: https://www.youtube.com/watch?v=T0vkiBPWigg&ab_channel=FilmScout.

¹³ The phenomenon of *emergence* constitutes a field of investigation in its own right. For the emergence of tools, see Morowitz, “Toolmaking”, 155-158. See also Fromm, *The Emergence of Complexity*. Bonabeau “Stigmergy.”

¹⁴ See Wittgenstein, *Philosophical Investigations / Philosophische Untersuchungen*, 39, §83: “And is there not also the case where we play and – make up the rules as we go along? And there is even one where we alter them – as we go along.” The corresponding German text (p. 39^e) reads: “Und gibt es nicht auch den Fall, wo wir spielen und – ‘make up the rules as we go along?’ Ja auch den, in welchem wir sie abändern – as we go along.”

vate the tool, in which case B will not have to face anything disagreeable, which B knows would befall him, if he would not act as agreed. This also holds for A. The use of the cooperation-enhancing tool, thus, represents a step away from the kind of practice depicted in Situation 2, which consisted in using the verb *agree* – in sentences like “This proposal *agree me*” – to make a statement about what was going on in B’s mind. In that case, the sentence can be understood as an *expression* of what B had already in his mind before he decided to utter it. B, thus, *informs* A about what he has in mind, or how he feels about the prospect of building a shelter with A. But in Situation 3, when B addresses A and utters the sentence “I *agree* with your proposal”, he is not informing A about anything; neither is he expressing an agreement that was already in their minds before he uttered the sentence. By addressing A and uttering the sentence “I *agree* with your proposal”, B empowers A to exact the relevant behaviour from B, if B does not act as agreed.

The cooperation-enhancing tool may have gradually emerged thanks to the influence of factors other than the careful consideration of reasons. But the consideration of reasons do still play a crucial role among the As and Bs. Seeing that tool worked as a powerful cooperation-enhancer, and that enhanced cooperation yields results that are agreeable to all of them, the As and Bs had good reasons to keep the tool and see what happens. I keep adding “and see what happens” as a reminder that the practice of making agreements, as any other human artefacts, is kept on the condition that its benefits consistently offset its costs. If the As and Bs see that the costs are too high, they have either of two options: they can abandon the practice of making agreements altogether or else keep the practice and amend it as they go along. If they decide to keep it and improve its cooperation-enhancing features as they further interact with one another, they may eventually become so dependent on the cooperation-enhancing tool that they will not even entertain the thought of throwing it away. As they “go along”, they will constantly propose amendments, but they will never abandon the basic practice of making agreements.

2. The Role of Sanctions

I have affirmed at the outset of this paper that contractarianism has been often criticised on the grounds of the assumption that we cannot derive norms from agreements unless we can firstly rely on some form of pre-agreement normativity. But the pre-agreement normativity, one might argue, cannot ultimately be accounted for on the grounds of yet another agreement. Somewhere the buck has to stop. But where? Do we simply have the insight that agreements are binding upon the agreeing

parties, or that *pacta sunt servanda*, as though this insight were grounded in a “law of nature” or a “categorical imperative of reason”?¹⁵ I have proposed that an agreement, at its most basic level, does not have to rely on some form of pre-agreement normativity because, ultimately, all that matters is the collective realization that something disagreeable will happen to A or B when one of them, after having activated the cooperation-enhancing tool, fails to act as agreed. No laws of nature, nor “categorical imperative of reason”, nor the assumption that persons are ends in themselves are required to understand why agreements are ultimately binding. Agreements ultimately require, as Peter Stemmer aptly puts it, a “system of sanctions.” In Stemmer’s account of moral contractarianism, sanctions play a crucial role. Consider for instance the following passage from his 2000 book-length account of moral contractarianism:

If a corresponding arrangement of interests and power obtains, one can reasonably achieve agreements about the establishment of moral norms through the creation of a system of sanctions. This approach is that of hypothetical contractualism.¹⁶

The account of agreements I have proposed above differs, though, from Stemmer’s in one important respect: how could the As and Bs possibly *agree* to the establishment of a system of sanctions before they had established what an agreement is? Previous generation of As and Bs may have established that when B activates the cooperation-enhancing tool, by addressing A and saying that he *agrees* to build a shelter with A, A will be entitled to impose something disagreeable on B, which B knows will befall him if B fails to act as agreed. But in doing so the As and Bs were not *agreeing* to establish a system of sanctions. It was, rather, the other way round: by engaging in a system of sanctions, the As and Bs were establishing what the act of *agreeing* is. They were thus changing the meaning words, as the verb *agree* came to mean now, primarily, what one commits oneself to, and not what one thinks is agreeable. It means

¹⁵ See Kant, *The Metaphysics of Morals*, 48: “Obligation is the necessity of a free action under a categorical imperative of reason.” For the original German text see Kant, *Die Metaphysik der Sitten*, 6:222: “Verbindlichkeit ist die Nothwendigkeit einer freien Handlung unter einem kategorischen Imperativ der Vernunft.”

¹⁶ Stemmer, *Handeln zugunsten anderer*, 205. The original German text reads: “Ist eine entsprechende Konstellation von Interessen und Macht gegeben, kommt es vernünftigerweise zu Agreements über die Etablierung moralischer Normen durch die Errichtung eines Sanktionensystems. Diese Vorgehensweise ist die des hypothetischen Kontraktualismus.” In this passage, the term “Kontraktualismus” might as well be translated as “contractarianism”.

that the As and Bs could not yet count on the institution of agreements in order to create a system of sanctions, or to change the meaning of words. They had first to engage in the new practice – which was also a new linguistic practice – and see what happens. Whether the new practice was invented for the sake of reasons or – as Wittgenstein put it – emerged as a sort of *game*, the rules of which are established and changed as the As and Bs “go along”, that is not the main focus of a philosophical inquiry, even considering that a philosophical inquiry has much to gain from other disciplines such as philology, archaeology, anthropology, evolution theory and so on. Given all we know about human beings, it is reasonable to assume that the new practice may have simply *emerged*. Be that as it may, the important thing is that the new practice consists in enhancing the prospect of cooperation by using a simple tool, namely: by addressing the relevant parties and uttering “I agree with your proposal. Let us do *this*.” If the As and Bs see that the new practice does indeed promote cooperation and that the cooperation is mutually advantageous, in spite of occasional defections, the As and Bs will opt for maintaining the practice and introduce amendments as they “go along.” And as they “go along”, the practice will become increasingly more complex, so complex that although new generations of As and Bs are constantly engaged in implementing amendments to new forms of cooperative arrangements, they correctly reason that they will never abandon the original practice of making agreements, whether or not, at its origins, the practice emerged itself from the consideration of reasons.

As they go along, the As and Bs may gradually see, for instance, that the pronoun *he* had hitherto been used to refer to As and Bs quite irrespectively of their being men or women; the As and Bs often spoke of *men*, when they actually meant human beings at large, as though only the male As and Bs were in a position to activate the cooperation-enhancing tool. The As and Bs may come to agree, then, that some aspects of the old linguistic practice must be amended. They (or some of them) tentatively decide to address this issue by using the pronoun *they* as a singular epicene pronoun, then they see what happens.¹⁷ The As and Bs who use other languages to communicate have to face a similar problem, as the female As and Bs grow increasingly dissatisfied with the current linguistic practice. They too, tentatively, start to implement amendments to old linguistic practices and see what happens. Thus, new generations of As and Bs may agree to change the agreements upon which their linguistic behaviour had been ruled, but they will not abolish the institution of agreement itself, for the practice of making agreements, at its most basic

¹⁷ The Economist, “English has a traditional solution to gender-neutral pronouns.”

level, has taken roots in their daily lives, shaping the way they behave towards each other and think about themselves. They will sometimes, for instance, experience a *disagreeable* feeling of guilt even while not having been found guilty by other people. They will also, sometimes, feel *disagreeable* feelings, which they call remorse and shame on occasions they believe they have failed to behave as agreed, even if nobody imposed any sanctions on them. At this point, the institution of agreements will have become so powerful in their daily lives and in the very image they have of themselves as human beings, that they now may be easily led into thinking that an agreement has a force of its own, quite regardless of the institutions in which agreements exist and emerged in the first place. And because now the new As and Bs hardly ever perceive the binding force of agreements as a force of their own making, they may feel inclined to believe that agreements (and related forms of commitment) bind them *a priori*, as though it were grounded in a “law of nature”, or a “categorical imperative of reason.” Consider for instance the following passage from Charles Larmore’s 1996 *The Morals of Modernity*:

We cannot view these commitments as ones we choose, or would choose, were we to stand back and reflect upon them. They are so integral to our very conception of ourselves as moral beings that to imagine them as objects of choice would be to imagine ourselves as without any guiding sense of morality – and so not only ill-equipped to actually choose them, but also lacking the right sort of identification with them. Such commitments seem better understood as felt convictions, the role of which is to shape whatever choices we do make. Their value is properly grasped when they are seen as the inherited basis of choice, not as the objects of choice.¹⁸

The account of agreements I have advanced thus far does not reject the assumption that some forms of commitments are “integral to our very conception of ourselves as moral beings.” On the contrary, it takes it as a given, and tries to make sense of that conception. It is, indeed, possible that the new As and Bs cannot even imagine a world devoid of the commitments they have toward each other, but we, philosophers, can. And the

¹⁸ Larmore, *The Morals of Modernity*, 130. See also Forst, ‘Moral Autonomy and the Autonomy of Morality: Toward a Theory of Normativity after Kant’, 58: “To paraphrase a famous remark by Heidegger on the ‘proof’ of the ‘external world’ one could say that it is not so much a scandal of philosophy that no answer has yet been found to the question concerning nonmoral interests in being moral that could convince the skeptic; rather, the scandal lies in the fact that an answer to this question is still being sought. For from the perspective of someone who understands himself as a moral being, from the perspective of moral ‘being-in-the-world’ so to speak, this question does not even arise; and someone who does not understand himself morally can never be brought to see the point of morality in this way. This is what is meant by the autonomy of morality.”

point of a philosophical investigation into the binding force of agreements is not to take agreements, or other forms of commitments, for granted, but, contrary to what Larmore suggests, “to stand back and reflect upon them.” I will return to this point in the last section of this paper.

3. The Problem of Legitimacy of Agreements

The account of how the As and Bs came to agree about building shelters and punishing defectors must face one important objection now. The objection is this: the individuals who are going to live under the sway of a system of sanctions have to agree with the establishment of the system of sanctions before the system is established in the first place, otherwise the system will not be *legitimate* to them. In other words, previous generations of As and Bs had to express their *consent*, either explicitly or implicitly, to the creation of a system of sanctions, if the system were to respect their *autonomy*. The account of agreements I have proposed above, apparently, cannot distinguish legitimate from illegitimate systems of sanctions. It seems to me, though, that this objection results from a misunderstanding.

If the As and Bs established the institution of agreements by engaging in a system of sanctions, they could not (strictly speaking) have agreed, or have given their consent, to the establishment of the system of sanctions. They had first to engage in the new practice and see what happens. If they realize that the cooperation-enhancing tool produces on average more harm than good, being too disagreeable to endure in the long run, they may prefer not to keep the tool. If the As and Bs, then, decide not to keep the tool, it is not because they have realized that the tool was illegitimate, nor because the tool represented a threat to their individual autonomy, but because they saw that, in the end, the tool did not enhance cooperation – it caused more harm than good. On the other hand, if they decide to keep the tool, because they see that on balance it does promote cooperation, they will also soon realize that the tool has some additional benefits: it enables them to build further (institutional) tools.

Once the As and Bs have mastered the practice of making agreements, the As and Bs will be in a position to speak not only of the things they have actually agreed to, but also of things they *would agree to* or *would not agree to*. New generations of As and Bs would probably not agree to the creation of new forms of cooperation that would rely on hanging people by the neck as forms of punishment for someone like B in Situation 3, simply because B failed to help A to build a shelter after having agreed to. They call a system of cooperation that they would not agree to create and support an *illegitimate* system. They can also speak of *legiti-*

macy and *illegitimacy* in order to refer to ongoing forms of cooperation. Thus, instead of saying that they are increasingly dissatisfied with the current linguistic practice, the female As and Bs might say that the current linguistic practice is not legitimate: they *would not agree* to create it because, in hindsight, they see now that the old linguistic practices has caused them more harm than good. From this it does not follow that the female As and Bs do actually have the power to change the current practice all at once. It only follows that their demand is *legitimate*.

If previous to the emergence of the original system of sanctions, by means of which the institution of agreements first came into existence, the As and Bs had to share an understanding about the legitimacy or illegitimacy of certain forms of human interaction, or about consent as a moral constraint on the pursuit of collective enterprises among autonomous human beings, then the critics of contractarianism would be right in saying that contractarianism must ultimately rely on some form of pre-agreement normativity related to ideas such as legitimacy, consent, and autonomy. But these ideas do not have to be taken for granted in the account of agreements I have proposed here. It does not follow from this, though, that concepts such as legitimacy, consent, and autonomy are devoid of meaning. On the contrary, these concepts constitute some of the most important conceptual tools we have created to regulate our interaction with each other. Yet, indispensable as they are, concepts such as legitimacy, consent, and autonomy must be ultimately comprehended as further developments of more basic forms of human interactions that unfold on the basis of agreements.

Shall we now agree that contractarianism does not have to rely on pre-agreement forms of normativity in order to derive norms from agreements? I fear some readers will remain unconvinced. They might argue that, in the end, I did not advance a good argument – or indeed, that I did not advance an argument at all. All I did was to tell a long fictional story about the successive generations of As and Bs who came to speak less and less about what was *agreeable* to them, speaking more and more about *agreements* among them. Why should a fictional story count as an argument in philosophy?

4. History, Hypothesis, or Fiction?

Admittedly, the As and Bs have never existed. The account on the emergence of agreements provided here is not meant to be understood as a piece of history, even though it has been informed by some historical records on the use of the verb *agree* in English and of some of its cognate forms in other languages. But then, if that is not a historical ac-

count, what kind of account is that? One might suggest, as Kant did in a commentary to Hobbes' theory on the state of nature, that the kind of account proposed here "is just an idea of reason" (*es ist eine bloße Idee der Vernunft*).¹⁹ John Rawls resumes Kant's understanding of the problem and suggests that the "original position" – Rawls' own theory on the state of nature – is a "hypothetical situation."²⁰ Each of the three situations I have depicted earlier might have been described as a "hypothetical situation" as well. With this move my account would be in line with a well-established tradition of social contract theories that have been often referred to as hypothetical contractualism or, as the case may be, hypothetical contractarianism.²¹ But if the situations depicted earlier are hypothetical in nature, rather than accounts of facts, how do they stand relative to other kinds of hypotheses? After all, scientific accounts can be presented in the form of hypotheses as well. Given all the evidence available, one might confidently affirm, for example, that the best hypothesis concerning the geographical origins of human beings is that they firstly emerged in Africa and then spread throughout the world. In the face of new evidence, we may possibly have to abandon this otherwise plausible hypothesis, but for now that is the best account for the geographical origins of humanity.²² Thus, the best account for the geographical origins of humanity that archaeology and anthropology provide is, indeed, a hypothesis, but that hypothesis purports to be more than a fictional narrative about our distant past. The narrative delivered by archaeology and anthropology purports to be true. The situations I have described earlier, on the other hand, do not purport to be true: they are fictional. There emerges the question, then, as to the role fictional narratives are supposed to play in a systematic enquiry into the concept of agreement.

Although the narratives depicted earlier do not purport to be true, they result from the application of a method that is not strange to other fields of inquiry.²³ Consider for instance how early modern scientists came to dis-

¹⁹ Kant, *Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis*, 8:297.

²⁰ Rawls, *A Theory of Justice*, 12-13: "In justice as fairness the original position of equality corresponds to the state of nature in the traditional theory of the social contract. This original position is not, of course, thought of as an actual historical state of affairs, much less as a primitive condition of culture. It is understood as a purely hypothetical situation characterized so as to lead to a certain conception of justice" [...] "Our social situation is just if it is such that by this sequence of hypothetical agreements we would have contracted into the general system of rules which defines it."

²¹ Gauthier, *Morals by Agreement*; Stemmer, *Handeln zugunsten anderer*; Scanlon, *What We Owe to Each Other*, 115, 176, 315; Moehler, "Contractarianism and climate change", 141; Williams, *Truth & Truthfulness*, 20-40.

²² Chan et al., "Human origins in a Southern African Palaeo-Wetland and first migrations."

²³ Meuter, "Narration in Various Disciplines."

cover some important features of the air. The air, quite literally, permeates our physical environment and is indispensable to human life. The air is to our physical environment what agreements are to our institutional environment – it is everywhere, and we could not go very far without either of them. Pervasive as it is, though, the problem with air is that we cannot really touch or see or smell or taste it in order to have a clear picture of its main properties. Most of the time we do not even take notice of its existence. One idea that occurred to some early modern scientists, then, was to create an environment devoid of air – and see what happens. Once they managed to build a vacuum pump, they saw that some strange things happened in the absence of air: they saw, among other things, that neither birds nor insects could fly, and that an alarm clock within a vacuum chamber could not be heard from the outside. They realized, thus, that the sound propagates through the air. The first systematic work describing experiments designed to understand the properties of air by creating an environment devoid of air was a book by Robert Boyle, published in 1660. Hobbes was well aware of these experiments.²⁴ Whether or not Hobbes had a similar methodological procedure in mind in his account of the state of nature, that is an interesting question, which I will not, though, try to pursue here, but the analogy I am driving at should be clear. If one wants to understand the nature of air, one has first to create an environment where air does not exist and see what happens. If one wants to understand which properties one specific gene (or group of genes) expresses in a given genome, one should silence it and see what happens. If one wants to understand the function of a line of code in a piece of software, one should delete it and see what happens. If one wants to understand which parts of the human brain are involved in our moral behaviour, one should examine a brain in which those structures do not exist, or have been damaged by an accident, and see what happens.²⁵ In like manner, if one wants to understand the nature of some of our most basic normative concepts, one has to create an environment where they do not exist – and see what happens. The difference is that we cannot (and should not) carry out this experiment literally. An institutional vacuum pump does not exist. We can only *imagine* how such an environment might look like, and how the persons living there would interact with one another – and see what happens. Would they try to fill the void if they had an opportunity to?

The three fictional situations depicted earlier are not useful illustrations for the purpose of an inquiry into the concept of agreement – they constitute

²⁴ Horstmann, *Leviathan und die Erpumper: Erinnerungen an Thomas Hobbes in der Luftpumpe*; Shapin and Schaffer, *Leviathan and the Air-Pump*; West, “Robert Boyle’s Landmark Book of 1660 with the First Experiments on Rarified Air.”

²⁵ Damasio, *Descartes’ Error*.

the very method that guides the inquiry itself. We literally imagine a narrative in which the concept under scrutiny – the concept of agreement – did not yet exist and ask if, starting from there, we would have reasons to arrive where we are, if we had an opportunity to. We imagine ourselves “standing back”, re-building everything from the scratch. If we can imagine that we would come up with something very similar, or as Stemmer puts the problem, something that “is in conformity” (*übereinstimmt*) with what we have now, then we would have reasons to keep the institutions we have, otherwise we should abandon, or at least try to amend them. As Stemmer put it:

The morality, as it really is, is either in conformity with the rational morality thus reconstructed or it is not. If it is in conformity, and is, therefore, rational, it is such that it might have arisen out of agreements, as simply imagined in the of hypothetical contractarianism approach. It is unimportant [*es ist gleichgültig*], though, how the morality actually emerged. What matters is not that it actually arose out of agreements, but that it is such that it could have emerged from agreements among rational people.²⁶

Fictional narratives, thus, play an important role for the purpose of a philosophical inquiry into the foundations of morality. They help us to establish a test for the legitimacy of our current social practices.²⁷ But contractarianism can only provide a reliable test for the legitimacy of our current practices if the fiction that guides the inquiry is at least informed by the best accounts available in other fields of knowledge such as philology, linguistics, archaeology, anthropology, evolution theory, ethnography and so on. Contrary to what Stemmer suggests in the passage above, I do not think that it is “indifferent” (*gleichgültig*) for the philosophical inquiry what other fields of investigation have to say about the factual origins of morality, or about the history of the words we use to convey normative ideas. Hypothetical contractarianism must be at least compatible with all we know about our evolutionary history, or with the history of human languages, and human institutions at large. Otherwise, the fictional account runs the risk of being just this – a piece of fiction.²⁸

²⁶ Stemmer, *Handeln zugunsten anderer*, 205-6. The original German text reads: “Die Moral, wie sie wirklich ist, stimmt mit der so konstruierten rationalen Moral entweder überein oder nicht überein. Wenn sie übereinstimmt und deshalb rational ist, ist sie so beschaffen, daß sie aus Agreements, wie sie in den Überlegungen des hypothetischen Kontraktualismus nur imaginiert werden, hervorgegangen sein könnte. Dabei ist es ganz gleichgültig, wie sie tatsächlich entstanden ist. Es kommt nicht darauf an, daß sie faktisch aus Agreements hervorgegangen ist, sondern darauf, daß sie so beschaffen ist, daß sie aus Agreements rationaler Personen entstanden sein könnte.”

²⁷ See Gauthier, “Why Contractarianism?”, 101: “Hypothetical agreement thus provides a test of the justifiability of our existent moral practices.”

²⁸ The author thanks Peter Stemmer for the institutional support and fruitful discussions in

References

- Araujo, Marcelo de. “Kontraktualismus”. In *Menschenrechte: Ein Interdisziplinäres Handbuch*, edited by Arnd Pollmann and Georg Lohmann, 193-98. Stuttgart: Verlag J.B. Metzler, 2012.
- Bonabeau, E. (1999). Editor’s Introduction: Stigmergy. *Artificial Life*, 5(2), 95–96. <https://doi.org/10.1162/106454699568692>
- Chan, Eva K. F., Axel Timmermann, Benedetta F. Baldi, Andy E. Moore, Ruth J. Lyons, Sun-Seon Lee, Anton M. F. Kalsbeek, et al. “Human Origins in a Southern African Palaeo-Wetland and First Migrations”. *Nature* 575, no. 7781 (7 November 2019): 185-89. <https://doi.org/10.1038/s41586-019-1714-1>.
- Chaucer, Geoffrey. *Troilus and Criseyde*. Translated by B. A. Windeatt. Oxford World’s Classics. Oxford; New York: Oxford University Press, 2008.
- Damasio, Antonio R. *Descartes’ Error: Emotion, Reason and the Human Brain*. New York: Quill, 2004.
- Forst, Rainer. ‘Moral Autonomy and the Autonomy of Morality: Toward a Theory of Normativity after Kant’. In *The Right to Justification: Elements of a Constructivist Theory of Justice*, translated by Jeffrey Flynn. New York: Columbia University Press 2012.
- Fromm, Jochen. *The Emergence of Complexity*. Kassel: Kassel University Press, 2004.
- Gauthier, David. *Morals by Agreement*. Oxford: Clarendon Press, 1986.
- . “Why Contractarianism?” In *Contractarianism, Contractualism*, edited by Stephen Darwall. Malden: Blackwell, 2003.
- Horstmann, Frank. *Leviathan und die Erpumper: Erinnerungen an Thomas Hobbes in der Luftpumpe*. Berlin: Mackensen, 2012.
- Jurafsky, Daniel. “Pragmatics and Computational Linguistics”. In *The Handbook of Pragmatics*, edited by Laurence R. Horn and Gregory Ward, 578-604. Oxford: Blackwell, 2006. <https://doi.org/10.1002/9780470756959.ch26>.
- Kant, Immanuel. *Die Metaphysik Der Sitten*. Akademieausgabe. Vol. 6. Berlin, 1793. <https://korpora.zim.uni-duisburg-essen.de/Kant/aa06/Inhalt6.html>.
- . *The Metaphysics of Morals*. Texts in German Philosophy. Cambridge: Cambridge University Press, 1991.
- . *Über Den Gemeinspruch: Das Mag in Der Theorie Richtig Sein, Taugt Aber Nicht Für Die Praxis*. Akademieausgabe. Vol. 8. Berlin, 1793. <https://korpora.zim.uni-duisburg-essen.de/Kant/aa08/273.html>.
- Kirkham, Richard L. *Theories of Truth: A Critical Introduction*. A Bradford Book. Cambridge: MIT Press, 1997.
- Klein, Ernest. “Agree (Intr. v)”. In *A Comprehensive Etymological Dictionary*

his research group at the University of Konstanz in 2021. Dario Teixeira, Pedro Fior Motta, and Daniel de Vasconcelos Costa also contributed with insightful comments. The author benefited from financial support provided by the Alexander-von-Humboldt Foundation (2021); FAPERJ (Research Support Foundation of the State of Rio de Janeiro, Brazil, Grant Nr. 202.643/2019); and CNPq (National Council for Scientific and Technological Development, Brazil, Grant No. 305050/2018-4). Last but not least, the author thanks two anonymous reviewers for their very insightful comments and suggestions.

- of the English Language, Dealing with the Origin of Words and Their Sense Development Thus Illustrating the History of Civilization and Culture*, Vol. 1. Amsterdam, London, New York: Elsevier, 1966.
- Larmore, Charles. *The Morals of Modernity*. Cambridge: Cambridge University Press, 1996.
- Mapel, David. "The Contractarian Tradition and International Ethics". In *Traditions of International Ethics*, edited by Terry Nardin and David Mapel, 180-200. Cambridge: Cambridge University Press, 2002.
- Meuter, Norbert. "Narration in Various Disciplines". In *Handbook of Narratology*, edited by Peter Hühn, John Pier, Wolf Schmid, and Jörg Schönert. Narratologia 242-262. Berlin, New York: Walter de Gruyter, 2009.
- Moehler, Michael. "Contractarianism and Climate Change". In *Moral Theory and Climate Change: Ethical Perspectives on a Warming Planet*, edited by Dale E Miller and Ben Eggleston, 139-56. New York and London: Routledge, 2020.
- Morowitz, H. J. (2004). *The Emergence of Everything: How the World Became Complex*. Oxford University Press.
- Prichard, H. A. (1949). Does moral philosophy rest on a mistake? In *Moral Obligation Essays and Lectures* (pp. 1–17). At Clarendon Press.
- Rawls, John. *A Theory of Justice*. Cambridge (Mass.): Harvard University Press, 1971.
- Reinach, Adolf. *Die apriorischen Grundlagen des bürgerlichen Rechts*, 1913.
- . *The a priori Foundations of the Civil Law. Along with the Lecture "Concerning Phenomenology"*. Translated by John Crosby. Texas: Aletheia, 1983.
- Scanlon, Thomas. *What We Owe to Each Other*. Cambridge (Mass.) and London: The Belknap Press of Harvard University Press, 2000.
- Shapin, Steven, and Simon Schaffer. *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. Princeton Paperbacks. Princeton: Princeton University Press, 1989.
- Skeat, Walter W. "Agree (v.)". In *An Etymological Dictionary of the English Language*, 11. Oxford: Clarendon Press, 1963.
- Stemmer, Peter. *Handeln zugunsten anderer: Eine Moralphilosophische Untersuchung*. Berlin; New York: W. de Gruyter, 2000.
- . "Moralische Rechte als soziale Artefakte". *Zeitschrift Für Philosophie* 50 (2002): 673-91.
- The Economist. "English Has a Traditional Solution to Gender-Neutral Pronouns". *The Economist*, 1 April 2017. <https://www.economist.com/books-and-arts/2017/03/31/english-has-a-traditional-solution-to-gender-neutral-pronouns>.
- West, John B. "Robert Boyle's Landmark Book of 1660 with the First Experiments on Rarified Air". *Journal of Applied Physiology* 98, no. 1 (January 2005): 31-39. <https://doi.org/10.1152/jappphysiol.00759.2004>.
- Williams, Bernard. *Truth & Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press, 2002.
- Wittgenstein, Ludwig. *Philosophical Investigations / Philosophische Untersuchungen*. Translated by G. E. M. Anscombe. Cambridge, Mass: Blackwell, 1999.