The Many Facets of Trust

Riccardo BARATELLA [a,1], Glenda AMARAL [b] Tiago Prince SALES [b]
Renata GUIZZARDI [b] and Giancarlo GUIZZARDI [b]

[a] *Free University of Bozen-Bolzano, Italy*
[b] *University of Twente, The Netherlands*

Abstract. Trust is an attitude that an agent (the trustor) has toward an entity (the trustee), such that the trustor counts upon the trustee to act in a way that is beneficial w.r.t. to the trustor's goals. The notion of trust is relevantly discussed both in information science and philosophy. Unfortunately, we still lack a satisfying account for this concept. The goal of this article is to contribute to filling this gap. First, we take issue with some central tenets shared by the main philosophical accounts, such as that there is just one relation of trust, that this relation has three argument places, and that trust is reliance plus some extra factor. Second, we provide a novel account of trust, also discussing different levels of trust. According to the account we put forth here, the logical form of trust sentences is expressed by a four-place relation. Further, we distinguish and characterize four kinds of trust relations and their connections. We also argue that trust and reliance are different phenomena. Third, on the basis of the proposed account, we extend the Reference Ontology of Trust (ROT). We call the new version of ROT that includes this extension "ROT 3.0". Finally, we discuss the implications of the new ontological definitions in the applications we have done of the concept of trust in other works, also pointing out future applications made possible by these novel accounts of trust.

Keywords. Trust, Belief, Reliance, ROT, UFO, OntoUML

1. Introduction

Human beings trust. We may almost push this claim so far to claim that it is plausibly essential to human beings that they trust. Trust is pivotal in everyday life: Claire trusts Tom to keep her secrets, and Tom trusts a *Anonymous Reviewer 2* to do a fair and precise revision of his article. Further, trust plays a fundamental role in information science, such as financial applications [1,2], requirement engineering [3], ethical systems [4], and enterprise architecture [5]. Finally, the notion of trust is extensively debated in philosophy – where several philosophical views provide different theories of trust [6].

Despite the paramount relevance of trust in everyday life, its role in information science and in philosophy, we still lack a satisfying account of trust. One of the main reasons, as we shall show, is that the existing accounts fail to acknowledge that there is not just one relation of trust, but several relations, in turn connected in interesting ways. We shall address this case of semantic overload of the term. Specifically, the goal of the paper is to characterize some crucial facets of trust relations that are particularly

important both in information science and in our everyday life. In particular, we shall distinguish and characterize four types of trust relations, and we shall show – contra the philosophical debates which takes trust to be reliance plus some extra factor [7,8] – that trust does not include reliance: they are different, albeit related phenomena.

In previous publications [9,2], some of us have proposed ROT, a Reference Ontology on Trust, grounded in the Unified Foundational Ontology (UFO) [10]. ROT answers questions about the nature of trust and the factors that influence it. In this work, we extend ROT to clarify and provide a deeper account of different types of trust relations and their connections.

Let us fix some minimal vocabulary to begin with. We shall call the cognitive agent who trusts "the trustor" and the entity trusted – not necessarily a cognitive agent – "the trustee". So, the trustor trusts the trustee. It is worth noticing two peculiar features of our account. First, we shall provide a non-motives-based account of trust. As argued by McLeod [6], trust accounts divide into motives-based and non-motives-based accounts, where the difference concerns the way in which the trustworthiness of the trustee is characterized. For there to be a trust relation, the trustee has to be trustworthy. Motives-based theories holds that the trustworthiness of the trustee is grounded in some specific motivations of the trustee. Non-motives-based account holds that the trustworthiness of the trustee is not grounded in specific motivations of the trustee. Second, our theory does not aim to provide a reductive account of trust – we are sympathetic to the idea that trust – and reliance – are primitive relations and we characterize them through a list of principles – not different from the way in which the notion of parthood is characterized in mereology [11].

The methodology we shall adopt is the following. First of all, we establish the logical form of trust claims; then, we proceed to characterize, at least partially, trust by employing a number of principles. We distinguish a core set of principles that are peculiar to what we shall call "Ground Trust", and additional principles that differentiate further other kinds of trust, and their connections.

The remainder of the article is structured as follows. Section 2 divides into three parts. In part one, we briefly chart the debate of trust in philosophy, highlighting some aspects or features that are objectionable in the current debate and that will be rejected in section 3. In part two, we briefly review UFO and OntoUML. In part three, we present ROT. Section 3 provides our novel conceptual analysis of trust. Section 4 builds upon this analysis to formulate a new and extended version of ROT, called "ROT3.0". Section 5 explores some implications of our novel account. Section 6 provides some final considerations.

2. Research Baseline

2.1. Charting Trust

Trust is an attitude that cognitive agents have towards entities in order to achieve certain objectives. According to the philosophical literature, the logical form of trust is a three-place relation "A trust B for C", where C is an action or a valued item [12,13]. Further, it is commonly held that there is just one notion of trust clarified by the following conditions: (i) the trustor is vulnerable – e.g., vulnerable to betrayal – and at risk – for

instance, the trustor risks that the trustee fails to do what the trustor is depending on them to do; (ii) the trustor relies on others to be competent to do what the trustor trusts them to do; (iii) the trustor relies on them to do it [6]. According to the philosophical literature, trust is reliance plus some extra factor, where the philosophical accounts strive to characterize what this something else is [7] [8]. Moreover, in order for there to be trust, the trustee must be trustworthy from the trustor's point of view. Philosophical accounts disagree about the nature of trustworthiness: what does it take for the trustee to be trustworthy? For sure, the trustee must be judged competent and motivated to pursue the goal of the trust relation by the trustor. But this clarification does not settle the what-is question concerning trustworthiness. The philosophical field divides in two groups. The first formation holds that the specific motivations of the trustee matter, and they account for the nature of trustworthiness in terms of these specific motivations. For this reason, this formation is called "motives-based theories". The second formation holds the nature of the motivations of the trustee does not matter. They account for the nature of trustworthiness in other terms. For this reason, the second formation is called "non-motives-based theories".

The initial clarification of trust brings with it several problems. First, while the logical form of trust sentences is captured by "A trust B for C", it seems that a trust relation also depends upon a certain context or situation. For instance, Tom trust Mary for helping him to find a job in the U.S., given the context that Tom has no Green Card and Mary is a member of the mob. However, Tom does not trust her to help him to find a job in U.S., in the context where Mary is not part of the mob. Thus, the issue is whether the logical form of trust claims must also include some parameter for contexts or situations. Furthermore, it is short of being plausible that there is just one kind of trust relation rather than many kinds of trust relations. Trust relations may be as diverse as the following ones: Renata trusts Brazil to win the 2026 FIFA World Cup; Tom trusts Mary to keep his secrets; Jim trusts Mary to do Section 5 of the paper; Jack trusts *Anonymous Reviewer 2* to do a precise and fair review. The third kind of trust relation involves a delegated task that is not present in the other relations; the fourth trust relation involves an explicit commitment of the trustee that may not be present in the other relations; the first kind of trust relation can happen even if the trustee is unaware of it and implies not commitments or delegations. Thus, it is natural to wonder whether there are different kinds of trust relation. Finally, the thesis that trust is reliance plus some extra factor seems implausible. Indeed, in the presence of reliance, the agent who relies can continue to look for evidence of reliability toward the entity they rely upon. However, if a trustor who trusts an entity y continues to look for evidence for the trustworthiness of the trustee y, then there is no trust at all. For instance, suppose that Tom trusts Mary for keeping his secrets, but he continues to look for evidence that she will keep them. Clearly, he is not trusting her. Therefore, trust and reliance come apart – contra the traditional literature. In this paper, we provide a conceptual analysis of trust that addresses these concerns, and that belongs to the non-motives-based theories formation.

2.2. The Unified Foundational Ontology (UFO)

The Unified Foundational Ontology (UFO) is an axiomatic domain-independent formal theory, developed by consistently putting together a number of theories originating from areas such as Formal Ontology in philosophy, cognitive science, linguistics, and philo-

sophical logic. UFO is divided into three incrementally layered compliance sets: UFO-A, an ontology of endurants (objects) [14], UFO-B, an ontology of events (perdurants) [15], and UFO-C, an ontology of social entities built on the top of UFO-A and UFO-B, which addresses terms related to the spheres of intentional and social things [16,17]. For an in-depth discussion and formalization, we refer the reader to [14,18,19,10,15].

In our proposal for extending the Reference Ontology of Trust, we mainly rely on some concepts defined in UFO-C. For this reason, in the remainder of this section, we focus our discussion on this ontology, briefly explaining a subset of its ontological distinctions that are relevant for our analysis.

In UFO-C, a social relator (e.g., a "service agreement") mediates a relation between two or more individuals (e.g., "Mary", and "Netflix") that play different roles in the relation (e.g., "service customer", and "service provider"). By participating in a social relator, the individuals bear a number of social commitments and claims. A social commitment is a commitment of an agent A towards another agent B. As an externally dependent moment [2], a social commitment inheres in A and is externally dependent on B. The social commitments necessarily cause the creation of an internal commitment in A. Also, associated to this internal commitment, a social claim of B towards A is created. Commitments and claims always form a pair that refers to a unique propositional content. For example, when Mary subscribes to Netflix, she commits (a social commitment towards that organization) to pay the subscription fee according to the chosen plan. Thus, it creates a social claim of Netflix towards Mary with respect to this particular propositional content (the payment). The social relator (the "service agreement"), besides relating the individuals ("Mary as service customer" and "Netflix as service provider"), is the mereological sum of all social commitments and claims that inhere in the corresponding individuals.

To conclude, UFO is formally connected to a conceptual modeling language (OntoUML) [14,20]. OntoUML was designed such that its modeling primitives (stereotyped classes and relations) reflect the ontological distinctions of its underlying ontology, and its grammar is enriched with semantically-motivated syntactical constrains that mirror UFO's axiomatization. UFO is also formally connected to a set of tools to facilitate the ontology engineering process, such as ontological design patterns and anti-patterns [21], visual model simulation [22], and transformations for codification technologies [23,24].

2.3. The Reference Ontology of Trust (ROT)

The Reference Ontology of Trust (ROT), is a UFO-based ontology that formally characterizes the concept of trust, clarifies the relation between trust and risk, and represents how risk emerges from trust relations [9,2]. It also provides a deep account of the different factors that can influence trust, such as mental biases, trustworthy behavior indications provided by the trustee and pieces of evidence that suggest a trustee should be trusted.

ROT (which fragment is depicted in Figure 1[3]), defines Trust as a complex mental state of a Trustor agent, composed of a set of Beliefs about a Trustee and its behavior.

---

[2]An externally dependent moment is an intrinsic moment that inheres in a single individual but that is existentially dependent on (possibly a multitude of) other individual external to its bearer (i.e., which is not the bearer's parts or intrinsic moments) [14].

[3]In all OntoUML diagrams, we adopt the following color coding: objects are represented in pink, qualities and modes in blue, and relators in green.

Trust is always about an Intention of the Trustor regarding a goal, for the achievement of which she counts upon the Trustee. The Trustor is necessarily an "intentional entity", that is, a cognitive agent, an agent endowed with goals and beliefs. As for the Trustee, it is an entity capable of impacting one's intentions by the outcome of its behavior [25], regardless of whether this involves an action or an omission (e.g. 'doing nothing', 'abstaining from doing X'). Further, an Intention is a Mental Moment (see also below), and every Intention is associated with a propositional content – termed "Goal". A Goal is an abstract representation of a class of situations that the agent wants to pursue. Situations in reality can satisfy the goal (the propositional content of an intention) [17]. We shall use expressions like, e.g., "achieve a goal" or "commit to a goal" in the sense of making true (or committing to making true) the propositional content of a certain intention.

In ROT, the trust mental state of a trustor regarding a trustee and its behavior is composed of: (i) an intention of the trustor; (ii) a set of beliefs about the trustee's capabilities and vulnerabilities; and (iii) if the trustee is an agent, beliefs that the trustee intends to exhibit the expected behavior. ROT relies on the notion of belief defined in UFO to model Aspectual Belief as a Belief that inheres in the Trustor and is externally dependent [14] on an (external) entity, the Aspect Type [18]. Aspect Type is a type[4] whose instances are Aspects. As defined in UFO-A, Aspect is a supertype of all modes, qualities and relators [18]. Dispositions, Intentions and Beliefs are examples of Aspects. In UFO, a belief is a special type of *mode*, named Mental Moment, which is existentially dependent on a particular Agent, being an inseparable part of its mental state. Other examples of Mental Moments are Intentions, Desires, Perceptions. As illustrated in figure 1, Aspectual Belief is specialized into Disposition Belief and Intention Belief. Disposition Belief is a belief about a Disposition[5] of the Trustee. It is distinguished into Capability Belief and Vulnerability Belief. The former refers to a belief about a Trustee's Capability — the Trustor believes that the Trustee is capable of performing a desired action or exhibiting a desired behavior — while the latter refers to a Trustee's Vulnerability — the Trustor believes that the Trustee's Vulnerabilities will not prevent it from performing the desired action or exhibiting the desired behavior. Capabilities and vulnerabilities are special types of dispositions. Capabilities are usually understood as positive dispositions, in the sense that they enable the manifestation of events desired by an agent. However, when the manifestation of a capability enables undesired events that threaten an agent's abilities to achieve a goal, it can be seen as a threat capability. Vulnerabilities are dispositions whose manifestation constitute a loss or can potentially cause a loss from the perspective of an agent. Note that the very same disposition may play the role of a capability, a vulnerability, or even a threat capability. For example, Bluetooth's functionality for transmitting and receiving data can be seen both as a capability and a vulnerability (as confidential information can be disclosed). As for the Intention Belief, it is a belief about an Intention of the Trustee — the Trustor believes that the Trustee intends to perform a desired action or exhibit an expected behavior.

Another ontological commitment made by ROT is that trust implies risk. By trusting, the Trustor becomes vulnerable to the Trustee in terms of potential failure of the ex-

---

[4]In UFO [14], types (or universals) can be seen as patterns of features that are repeatable across individuals Individuals instantiate types.

[5]In UFO [15], *dispositions* are dependent entities in the sense that they depend on other objects to exist. They are considered as moments that are only manifested in particular situations on the occurrence of certain triggering events, and that can also fail to be manifested.
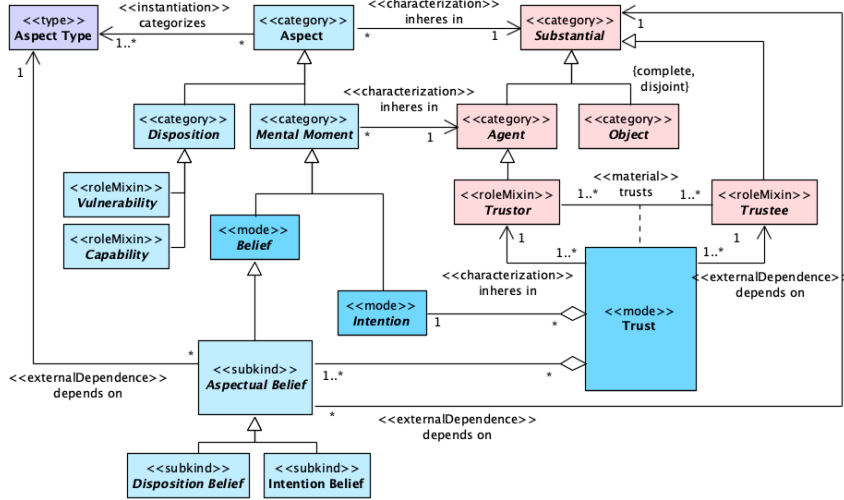
Figure 1. A fragment of the Reference Ontology of Trust (ROT).

pected behaviour or outcome [26, p. 21]. In trust relations, risk can emerge (as threat and loss events) [27] as a consequence of either the manifestation of a Trustee's Vulnerability or the unsatisfactory manifestation of a Trustee's Capability.

3. The Many Flavors of Trust and Betrayal

3.1. Ground Trust

Let us first fix the logical form of trust sentences by considering the following ones: (1) John trusts Michael to help him to find a job in US; (2) Tom trusts Federer to play tennis well during the Wimbledon Tournament.

Sentences 1-2 suggest that the logical form of a trust sentence involves a four-place relation displayed as follows: **(LF)** *x* **trusts** *y* **for** *g* **in situation** *s*, where *x* is the trustor, *y* is the trustee, *g* is the goal of the trustor for the achievement of which the trustor counts upon the trustee, situation *s* is the specific situation of the trust relation and is an instance of a certain situation type *S*. Specifically, in this work, we conceive a situation as a partial possible world (in the sense of [28]) that must be compatible with goal *g*, and that provides the background conditions of the trust relation.

LF makes it clear a sharp difference between the suggested account and the current theories on the market, that take a trust relation to be a three-place relation "*A* trusts *B* to do *X*" [12,13]. The reason for our revision of the current literature is based on the following fact. Consider sentence 2. Tom trusts Federer to play tennis well during the Wimbledon Tournament, because Tom has justified beliefs that Federer is a great tennis player on grass surface, but Tom does not trust Federer to play tennis well during the Roland Garros Tournament because Federer's mode of play does not fit well with clay surface. In this case, *playing on grass* and *playing on clay* are just different CAPABILITIES, which as dispositions are activated in different ways in different situation types. So, it is clear that the specific type of situation matters in order for there to be trust – contra the traditional

lore in philosophical debates. As we shall suggest below, it is a specific situation type that fixes, at least partially, to have one trust relation instead of another. Second, our initial clarification of *g* as the goal of the trustor for the achievement of which the trustor counts upon the trustee makes it clear that trust does not entail delegation. Consider sentences 1 and 2 again. While there is a plausible delegation in sentence 1, this delegation feature is completely absent in 2. So, delegation is not a necessary feature of any trust relation.

With the logical form LF of trust relations in place, let us distinguish different relevant kinds of trust relation. We shall start by fixing some necessary principles that characterize what we call "Ground Trust". To be clear, we don't take that these principles provide a definition of Ground Trust, nor are they necessary and jointly sufficient conditions to have trust, i.e., we admit that this list may not be complete. Here are the first six principles adapted from the original version of ROT to characterize **Ground Trust**:

1. The trustor has an intention that concerns a goal the trustor wants to achieve and for which she counts on the trustee. For instance, in a windy day, Tom has the intention to hold his papers in place, and he counts upon his laptop to hold his papers in place (by acting as a paperweight). In this case, the fact that the day is windy is part of the situation.

2. The trustor believes that the trustee has the capabilities to perform the relevant act for the goal to be achieved. For instance, Tom believes that his laptop has the capabilities to pursue the task, namely has a weight and a shape that will hold the papers in the place.

3. The trustor believes that the vulnerability of the trustee will not prevent her from performing the task. For instance, Tom believes that his laptop is vulnerable to the weathering. But Tom believes that this vulnerability won't prevent his laptop to keep his papers safe from the wind.

4. The trustor is necessarily an intentional agent. In other words, the trustor is an agent endowed with goals and beliefs – call it "a cognitive agent".

5. The trustee is not necessarily an intentional agent. The trustee may just be an entity, like an iPhone, that is capable of having an impact on a goal of the trustor.

6. Trust always implies risk and vulnerability. By trusting, the trustor accepts to become vulnerable to the trustee in terms of potential failure of the expected action and result, as the trustee may not perform the expected action or the action may not have the desired result. For instance, Tom is vulnerable since his laptop can fail to be a good paperweight in case of very strong wind.

A further principle that we suggest to characterize Ground Trust is the following:

7. If the trustee is a cognitive agent, then the trustor believes that the trustee has the motivation to act toward the goal of the trust relation. For instance, Tom believes that Federer is motivated to play tennis well during the Wimbledon Tournament. It is important to notice that the motivations of the trustee may be different as they can be – however, to the extent that these motivations allow the trustor to achieve their goal, they work. This is the reason why the account suggested is a non-motives-based account of trust – we do not hold that the motives for the trustee to act must be of a specific kind, such as self-interested [12], goodwill [29] or virtue [30]. In particular, a murderer may trust another murderer to keep their secrets – and this trust relation is not based on the good-will or the virtue of the

murderer-trustee: the murderer-trustee's reasons to act are based on his objective not to bother another killer.

Principle 7 is conditional to the fact that the trustee is a cognitive agent. So, it is not necessary for every ground trust relation. However, if the trustee is a cognitive agent, then principle 7 holds. Thus, strictly speaking, there are two kinds of Ground Trust relation: the first one has it that the trustee is a cognitive agent and principle 7 holds, and the second one covers the cases in which the trustee is not a cognitive agent and so principle 7 doesn't hold. Otherwise stated, for the ease of presentation, we shall continue to include Principle 7 in Ground Trust.

Now, we claim that principles 2-3-7 together provide a minimal ground for the trustor's belief of the trustworthiness of the trustee. There cannot be trust without the trustor's belief of *trustworthiness* of the trustee. The suggested account claims that the fact that the trustor believes that the trustee is trustworthy is grounded (in the philosophical sense – [31]) in 2-3-7. However, the ground for the trustor's belief of trustworthiness of the trustee is not exhausted by 2-3-7. Instead, it may include also evidence not included in the previous principles – e.g., the evidence that Federer played tennis well during the previous editions of the Wimbledon Tournament – as well as trustor's biases – as in the case where Tom trusts Lucas because they both support the same football team.[6] Further, these principles make it clear that the suggested account doesn't adopt an objective standard of trustworthiness, but it gives an internalist account of trustworthiness – trustworthiness from the perspective of the trustor. It is also worth noticing that, since the principles focus on the trustor's beliefs, these beliefs – including the one concerning the trustworthiness of the trustee – may well be false.

With these principles in place, let us consider a further crucial principle:

8. Trust requires that the evidence supporting trustworthiness cannot be complete.

To show the plausibility of 8, consider the following example (inspired by [6]). Suppose that Jenna and Jim want to have complete evidence toward their new babysitter Betty. Then, there is a clear sense in which Jenna and Jim don't trust Betty as a babysitter.

These conditions characterize, at least partially, Ground Trust, namely the weakest possible form of trust. Examples of Ground Trust include: Tom trusts Federer to play tennis well during the Wimbledon Tournament, Tom trusts his laptop to hold his papers in the place, and so on. We also suggest that an agent may trust an abstract entity or an abstract system – such as quantum mechanics – for a certain goal in a certain situation. Given the framework provided, trust in abstract entities – e.g., Trust in Humanities meaning in the qualities that should make humans humans – is a form of ground trust.

Ground Trust has not been recognized in the philosophical literature [6], either because it is not a motives-based account or because it doesn't include any implicit or explicit commitment of the trustee toward the trustor. However, the discussed everyday examples show that it is crucial to distinguish a relation of Ground Trust, where no commitment is done, from types of trust where this commitment is done. It is to these trust relations that we now turn our investigation.

---

[6]It is worth noticing that we propose a characterization of trust according to which trust comes in degree, rather than being completely present or not.

3.2. Weak and Strong Trust

The other forms of trust build upon Ground Trust by adding principles to those previously listed. A crucial pair of trust relations involves the supposed commitment of the trustee towards the trustor's goal. We have two types of trust identified by the following principles, respectively.

Weak Trust. The principle to be added to have Weak Trust is principle 9: the trustor has the belief that the trustee has a commitment to *goal c*, which the trust relation is about. For instance, Tom trusts his friend Lora to tell him the truth. It is because Tom and Lora are friends that he believes she is committed in accomplishing *goal c* that is the goal upon which this trust relation is founded. However, this belief may very well be false. It is for this reason that trust is so often associated with the notion of betrayal (something we examine in the sequel). It is crucial to note the notion of commitment referred here is one of *social commitment*, which entails social claims concerning *goal c* on part of the trustor. Further, when the trustee is not a cognitive agent, we shall adopt a kind of indirect commitment, e.g., the commitments of a car or a clock are the commitments of the producer of these items (see below for examples).

Strong Trust. Strong Trust is obtained by adding principle 10 to those listed for Ground Trust: the trustor knows that the trustee explicitly commits himself to *goal c*, which the trust relation is about. For instance, suppose that Tom trusts *Anonymous Reviewer 2* to review his article in a fair and careful way. Then, this trust relation presupposes the explicit social commitment of *Anonymous Reviewer 2* to review Tom's article, and Tom knows about this fact. At least in the case of Strong Trust, it seems that the explicit commitment of the trustee has consequences also for the commitments, the beliefs, and the acts of the trustor. For instance, if Mary trusts John to cook dinner and John explicitly commits himself to it, then Mary will act as if she won't cook the dinner, she is committed not to cook the dinner, and she believes she will not cook the dinner. What is more, if Mary started cooking the dinner, then John may rightly be frustrated, because at least he has the belief that Mary won't cook the dinner. Thus, at least in presence of 10, the commitment of the trustee seems to generate some commitment on the side of the trustor too. Clearly, 9 and 10 are not incompatible requirements. Suppose that the trustee explicitly commits himself to *goal c* the trust relation is about, and the trustor knows about this commitment. Then, principle 9 follows, given the premise that knowledge implies belief.

3.3. Trusted Delegation

Now, the previous kinds of trust relation didn't take into consideration the phenomenon of delegation. Many trust relations involve the fact that the trustor delegates *goal c* of the trust relation to the trustee. For instance, suppose that the trustor is the main author of the paper and trusts one of their co-authors to do a part of the paper either because of the lack of time or because the co-author is more versed in this respect. We now turn to the relation between trust and delegation.

Webster dictionary defines delegation as "the act of empowering to act for another". This is consistent with the concept of delegation adopted here, which is the one argued in [16]: a material relation between a delegator and a delegatee, which involves a social relator composed of pairs of commitment/claim and dependence between delegator and

delegatee. So defined, delegation clearly does not entail trust. There may be delegation without trust, due to coercion or because one does not have another option but to delegate. For example, suppose one's car breaks down in the middle of nowhere. After a few hours, a driver stops and says he can fix it for 100 dollars, but he has to go grab his tools in another town first and he demands to be paid in advance. In this case, the car owner indeed delegates (thus, creating corresponding commitments and claims) the goal of having the car fixed to the alleged mechanic. The car owner does that because he has no other choice, even if he does not trust the delegatee.

We define a special type of delegation relation grounded on trust that we will term **Trusted Delegation** 11, i.e., a delegation relation in which the trustor delegates to the trustee to achieve *goal c*, i.e. the goal that the trust relation is about.

While Weak Trust is Ground Trust plus 9, and Strong Trust is Ground Trust plus 10, the challenging question is what kind of connection there is between Trusted Delegation and different forms of trust. Clearly, Trusted Delegation is compatible with both Weak Trust and Strong Trust – namely, Weak Trust and 11 as well as Strong Trust plus 11 are consistent. At minimum, if the trustor deliberately delegates to the trustee to achieve *goal c* upon which the trust relation is founded, then the trustor must have the belief that the trustee is committed to achieve *goal c*. However, it is too strong to require that Trusted Delegation entails Strong Trust. For instance, John trusts his wife to go get their children at School, so he delegates this task to her. However, she might not have explicitly committed herself to this goal. Hence, Trusted Delegation requires at least Weak Trust. In other words, for Trusted Delegation to exist, there must be either Weak Trust or Strong Trust.

3.4. Frustration and Betrayal

Let us examine the connection between trust and betrayal. When a trust relation fails, the trustor feels frustrated and betrayed. But why is it so? Given the suggested analysis, a necessary principle characterizing Ground Trust is 6: "Trust always implies risk and vulnerability." – namely, by trusting, the trustor accepts to become vulnerable to the trustee in terms of potential failure of the expected action or behavior.

We may individuate different forms of frustration or betrayal corresponding to the different kinds of trust. First, the ground form of betrayal, that we call "frustration", corresponds to the failure of Ground Trust. For instance, this situation happens when Tom trusts Federer to win the Wimbledon Tournament, but Federer loses. In this case, the trust relation fails because of the failure of Federer's expected action. As a result, the vulnerability associated to the trust relation is manifested in the form of *frustration*. We cannot say that Tom has been betrayed in this case, since there was no commitment from Federer to Tom (and even no plausible expectation of implicit commitment) regarding the goal of winning Wimbledon.

Let us now examine what happens in the case of Weak, Strong, and Trusted Delegation, respectively. For instance, when Tom believes that Mary commits herself to *goal c* the trust relation is about, we have Weak Trust. For example, because Mary is Tom's friend, he has the plausible expectation of commitment of Mary not to disclose secretes revealed to her. If Mary breaks Tom's believed commitment and Tom knows it, the trust relation fails and there is a kind of betrayal that we call ***Weak Betrayal***. Weak betrayal is what happens in all cases where there is an expectation of commitment without an

explicit one. This typically includes the cases in which the expectation is entailed by the role a trustee plays in a given relation to the trustor. Another example is when a political candidate from a Labor Party, once in power, creates legislation that hurts the interest of workers. Their electorate would be warranted to feel (weakly) betrayed but not in the same way as if he breaks explicit commitments made during the campaign. The latter stronger form of betrayal is examined in the sequel.

In case of Strong Trust, Mary explicitly gives her commitment - imagine Tom saying to Mary: "please do not tell my secret to anyone", to which she replies "no problem, I won't.". Thus, if she fails to honor this commitment, there is a kind of betrayal that we call ***Strong betrayal***. Betrayal and knowledge of betrayal interact in non-trivial manners. For example, if Tom knows about Mary's (weak or strong) betrayal, he will probably feel betrayed. But, in the case of an explicit commitment, even if Tom doesn't know about Mary's failing to honor her commitment, the (strong) betrayal still exists. In the latter case, since Tom doesn't know about it, he might still trust her, i.e., the trust relation might still hold. Further, there may be the failure of a trust relation because the trustor has the belief that the trustee broke their explicit or believed commitment – even in the case this commitment is still in place. If this is the case, then there is a "Weak Betrayal Belief" or a "Strong Betrayal Belief", corresponding to Weak and Strong Trust, respectively.

Finally, let us consider Trusted Delegation, which being based on principle 11, entails either Weak Trust or Strong Trust. If the trustee fails to achieve the delegated goal of the trust relation, then we call this ***Delegation Betrayal***. A Delegation Betrayal adds up to either Weak Betrayal or Strong Betrayal, depending on whether the commitment is just believed by the trustor or explicitly claimed by the trustee.

In summary, our analysis shows that the notion of trust has many facets, i.e., there are different kinds of trust relations. The natural issue is when there is one relation of trust rather than another one. A not-so-much interesting answer is "go and check which principles are present in a given situation". However, this not-so-much interesting answer suggests another, interesting principled answer: it is a given situation type together with a given goal that require one trust relation rather than another. Consider the situation types "asking for help to find a job in the US given a specific context" and "telling secrets to a person". The first situation type requires that a trustor trusts someone to help her find a job in the US given a specific context. In turn, this trust relation entails, at least partially, a delegation of the goal to find a job in the US. Thus, the situation type at issue demands a Trusted Delegation. The second situation type requires that the trustor trusts someone to keep her secrets. It also plausibly requires that the trustor at least believes that the trustee is committed to that goal. This trust relation is, at least, a kind of Weak Trust. Thus, the second situation type demands, at least, Weak Trust. Hence, it is plausible that a certain situation type together with a given goal require certain trust relations rather than others.

To conclude our account of trust relations, we briefly discuss the relations between trust and *reliance*. The current philosophical literature holds that trust is reliance plus some extra factor [8,7,6]. This seems not plausible. As the previous analysis showed, a trust relation necessarily requires that the evidence supporting trustworthiness cannot be complete. Further, our analysis showed that we can have Ground Trust without having stronger forms of trust – in particular, Weak Trust or Strong Trust. However, the ordinary notion of reliance is the one according to which I rely on my clock to ring in the morning or I rely on Amazon to deliver certain goods. On the one hand, reliance is compatible with – ideally – the complete evidence supporting the reliability of what I rely upon. One

may continue to ask for evidence for the reliability of one's clock to ring, or Amazon to deliver goods without failing to rely on them. On the other hand, it seems that if x relies on y for c in situation d, the relation of reliance requires either that the reliant x believes that the relied-upon y committed to the *goal c*, or that the reliant x knows that relied-upon y explicitly committed to *goal c*.[7] In conclusion, a result of our analysis is that trust and reliance concern different phenomena. In particular, reliance can be seen as a limiting case of either Weak Trust or Strong Trust.

4. The Proposed Model (ROT 3.0)

In this section, we extend the Reference Ontology of Trust (ROT) to formalize the assumptions on the different types of trust relations and their connections, discussed in the previous section. We present the new version of ROT that includes these extensions, namely ROT 3.0, in the OntoUML model depicted in figure 2.

In ROT 3.0, the previous entity Trust (figure 1) is renamed Ground Trust (figure 2), as it corresponds to the concept of ground trust described in Section 3.

ROT 3.0 introduces the concept of Weak Trust, which is a specialization of Social Trust, in which the Trustor has the belief that the Trustee has a commitment to the goal the Trust relation is about (that is, the propositional content of the Trustor's Intention). Social Trust is a specialization of Trust in which the Trustee is an Agent [9,2]. The Intention Belief is specific to this type of trust, as in this case the Trustee is a cognitive agent endowed with goals. We model trustor and trustee's commitments (entities Trustor Commitment and Trustee Commitment, respectively) as a Social Commitments (see section 2.2) [16,17]. The Social Commitment Belief represents the Trustor's belief that the Trustee is committed to the goal, for the achievement of which she counts upon the Trustee. It is modeled as a Belief that inheres in the Trustor and that is externally dependent on Aspect Type. According to UFO-C [16,17], Social Commitments are types of Aspects, and therefore instantiate Aspect Type (see figure 1).

Strong Trust is another type of trust relation proposed in this extended version of ROT, in which the Trustor knows that the Trustee explicitly commits herself to the goal the trust relation is about. Here we also apply the concept of social relator (see section 2.2) to model Strong Trust as a social relator that is grounded on Weak Trust, which is composed of pairs of social commitments and social claims that inhere either in Trustee or in the Trustor, as explained is section 3. Briefly speaking, Strong Trust mediates the relation between the Trustor and Trustee by being a social relator aggregating their explicit commitments and claims.

Finally, as argued in the previous sections, many trust relations involve the fact that the trustor delegates the goal of the trust relation to the trustee. To account for this, ROT 3.0 introduces the concept of Trusted Delegation. We model Trusted Delegation as a social relator, grounded on Weak Trust, which mediates a delegation relation between the Trustor (the delegator) and the Trustee (the delegatee).

---

[7]If the relied-upon entity is something like a clock, the commitment is indirect. See Section 3.2 on Weak Trust.
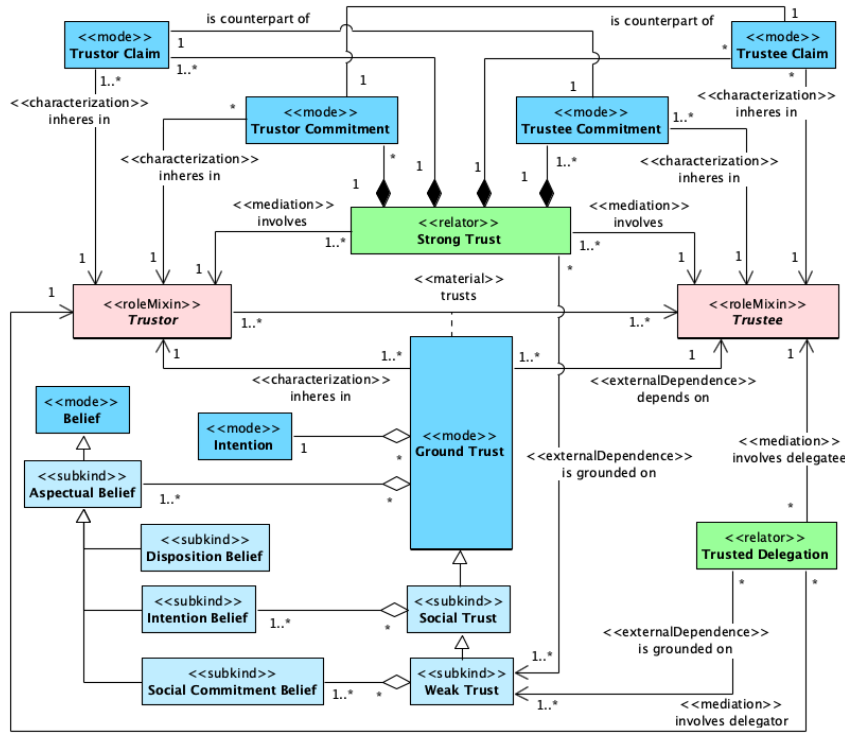
Figure 2. A fragment of ROT 3.0 depicting Weak Trust, Strong Trust and Trusted Delegation.

5. Implications

In recent works, we have explored the ontological notion of trust defined in ROT, namely in analyzing trust in Enterprise Architecture by proposing a pattern language for Archi-Mate [5], and in systematically eliciting and representing trustworthiness requirements to support the development of trustworthy Information Systems [3,1]. Now, we must analyze the implications of the definitions of weak trust, strong trust and trusted delegation to these works. Since these distinctions were not present, and we had only one concept of trust (in ROT 3.0, understood as ground trust), then it is clear that in those works, we collapsed every definition related to the pattern language and to trustworthiness requirements on a sole trust concept.

Let us take an example of trustworthy systems. When we say that Anna trusts her autonomous car to drive her to work safely, is this an example of *ground*, *weak* or *strong trust*? Moreover, is this a case of *trusted delegation*? In this example, Anna is the trustor, the autonomous car is the trustee and the trusted goal is "to drive her to work safely". This is a typical case of trust, which entails risk and vulnerability towards the trustor. But in order to define which kind of trust this is, we must be sure to understand if this case includes only beliefs about capacities and intentions, also beliefs about commitments, or actual commitments. Anna thinks her autonomous car is indirectly committed to drive her to work safely (Commitment Belief); but is the car explicitly (indirectly) committed? If we consider the contract between Anna and the autonomous car producer, yes, and this would be a case of *Strong Trust*. Moreover, there is a dependence between Anna and

the autonomous car, without which she is not arriving at her work. In other words, she delegates to her car to take her to work safely. Thus, we affirm that this is a case of *Trusted Delegation*. We may not however take for granted that all cases of trust in Information Systems are cases of Strong Trust. This needs to be analyzed according to the existence or not of commitments and claims from both sides. It is not possible to claim that the trust we have on a software system to perform operations for which it wasn't designed (i.e., *software repurposing*) to be a case of Strong Trust. Ground Trust seems to be still the best way to describe such relation. By considering all these finer-grained distinctions, we can evolve the requirements engineering approach proposed in [3,1].

Similarly to the case of trustworthiness requirements, the ArchiMate pattern language of trust needs to be revisited to verify how the different kinds of trust appear in that work. Moreover, we envision the potential for new applications, for example, to analyze different kinds of services contract for manual or automated services. Trust has a great impact on how people acquire and use services. Additionally, trust also influences the kinds of claims clients have when they suffer losses caused by poor services. So this may be an interesting avenue to be explored.

6. Conclusions and Future Perspective

In this article, we provided a non-motives-based and internalist account of trust – where an internalist account of trust means that trust is conceived from the point of view of the trustor. We claimed that the expression "Trust" is an umbrella term for different kinds of relations and we distinguished and characterized four kinds of trust relations, laying down their connections. We also suggested that different kinds of trust relation lead to different kinds of betrayal, and we distinguished three of them, besides the concept of frustration. Moreover, we argued that trust and reliance are different phenomena. On the basis of the proposed account, this paper presented an OntoUML model extending the reference ontology ROT based on the new definitions we introduced. This OntoUML model makes it clear that Trust is a systematically polysemic term and that the underlying ontological notion is a cross-categorical notion: Ground Trust and, hence, Weak Trust are externally dependent modes that induce corresponding *Cambridge relations* of trust, i.e., relations that do not bestow any new aspects to the trustee; in contrast, Strong Trust is a relator, bestowing commitments and claims to both participants and grounding a genuine material relation of trust. Finally, we also discussed some implications of the given account in recent works, besides new applications.

As future works, we plan to examine the issue of whether and under what conditions trust relations are justified and unjustified, and the topic of the biases of the trustor and its implications for trust relations. We also plan to investigate how trust propagates among several subjects, what links of dependence this phenomenon creates, and how it differs from risk propagation. Applying these distinctions in application scenarios discussed in Section 5 is also part of our research agenda.

Acknowledgement

References

[1] Amaral G, et al. Trustworthiness requirements: The Pix case study. In: Conceptual Modeling. ER 2021. vol. 13011. Springer; 2021. p. 257-67.

[2] Amaral G, et al. Ontological foundations for trust dynamics: The case of central bank digital currency ecosystems. In: Research Challenges in Information Science. RCIS 2022. Springer; 2022. p. 354-71.

[3] Amaral G, et al. Ontology-based Modeling and Analysis of Trustworthiness Requirements: Preliminary Results. In: Conceptual Modeling. ER 2020. vol. 12400. Springer; 2020. p. 342-52.

[4] Guizzardi R, et al. Ethical Requirements for AI Systems. In: Advances in Artificial Intelligence. vol. 12109. Springer; 2020. p. 251-6.

[5] Amaral G, et al. Modeling trust in enterprise architecture: A pattern language for ArchiMate. In: The Practice of Enterprise Modeling. PoEM 2020. vol. 400. Springer; 2020. .

[6] McLeod C. Trust. Stanford encyclopedia of philosophy. 2006.

[7] Goldberg SC. Trust and reliance 1. The routledge handbook of trust and philosophy. 2020:97-108.

[8] Hawley K. Trust, distrust and commitment. Noûs. 2014;48(1):1-20.

[9] Amaral G, et al. Towards a reference ontology of trust. In: On the Move to Meaningful Internet Systems. OTM 2019 Conferences. vol. 11877. Springer; 2019. p. 3-21.

[10] Guizzardi, G et al. UFO: Unified foundational ontology. Applied Ontology. 2022;17(1):167-210.

[11] Cotnoir AJ, Varzi AC. Mereology. Oxford University Press; 2021.

[12] Hardin R. Trust and trustworthiness. Russell Sage Foundation; 2002.

[13] Baier A. Trust and antitrust. Ethics. 1986;96(2):231-60.

[14] Guizzardi G. Ontological foundations for structural conceptual models. The Netherlands: Telematica Instituut, University of Twente; 2005.

[15] Benevides AB, et al. Representing a reference foundational ontology of events in SROIQ. Applied Ontology. 2019;14(3):293-334.

[16] Guizzardi RSS, Guizzardi G. Ontology-based transformation framework from TROPOS to AORML. In: Social modeling for requirements engineering. The MIT Press; 2010. p. 547-70.

[17] Guizzardi G, et al. Grounding software domain ontologies in the Unified Foundational Ontology (UFO). In: 11th Ibero-American Conference on Software Engineering; 2008. p. 127-40.

[18] Guizzardi, G et al. Types and taxonomic structures in conceptual modeling: A novel ontological theory and engineering support. Data & Knowledge Engineering. 2021;134:101891.

[19] Fonseca C, et al. Relations in Ontology-Driven Conceptual Modeling. In: Conceptual Modeling. ER 2019. vol. 11788. Springer; 2019. p. 28-42.

[20] Guizzardi G, et al. Towards ontological foundations for conceptual modeling: the Unified Foundational Ontology (UFO) story. Applied ontology. 2015;10(3-4):259-71.

[21] Ruy, F et al. From reference ontologies to ontology patterns and back. Data & Knowledge Engineering. 2017;109:41-69.

[22] Benevides AB, et al. Validating Modal Aspects of OntoUML Conceptual Models Using Automatically Generated Visual World Structures. J Univers Comput Sci. 2010;16(20):2904-33.

[23] Almeida JPA, et al. gUFO: A Lightweight Implementation of the Unified Foundational Ontology (UFO). online: `https://purl.org/nemo/doc/gufo`; 2020.

[24] Guidoni, G et al. Transformation of ontology-based conceptual models into relational schemas. In: Conceptual Modeling. ER 2020. Springer; 2020. p. 315-30.

[25] Castelfranchi C, Falcone R. Trust theory: A socio-cognitive and computational model. vol. 18. John Wiley & Sons; 2010.

[26] Luhmann N. Trust and power. John Wiley & Sons; 2018.

[27] Sales TP, et al. The Common Ontology of Value and Risk. In: Conceptual Modeling. ER 2018. vol. 11157. Springer; 2018. p. 121-35.

[28] Kratzer A. Situations in natural language semantics. Stanford encyclopedia of philosophy. 2007.

[29] Jones K. Second-hand moral knowledge. The Journal of Philosophy. 1999;96(2):55-78.

[30] Potter NN. How can I be trusted? A virtue theory of trustworthiness. Rowman & Littlefield; 2002.

[31] Correia F, Schnieder B. Metaphysical grounding: Understanding the structure of reality. CUP; 2012.