Trust, choice, and self-preservation: a computational approach

Piotr Cofta[1]

Abstract

Trust tends to be described through the lens of a rational choice of a trustor driven by the trustworthiness of a trustee. This, however, does not exhaust scenarios where people seem to be comfortable trusting without having an actual choice or not trusting while having a selection of potential trustees. This paper proposes that there should be another force at work that only expresses itself through the rationality in situations of choice. It is the self-preservation that strongly influences our decisions, specifically when choices seem to diminish. Ignoring this force makes the theory of trust incomplete. This paper brings this area of social behaviour closer to computational trust by proposing a unifying model that builds on theories of social systems. It focuses specifically on the containment of complexity and the associated risk to self-preservation, where trust is not an option but a necessity. The model, by being both simple and expressive, can computationally explain several phenomena associated with trusting in situations where self-preservation may be under a threat. This is further demonstrated by several use cases.

Keywords Computational trust · Social systems theory · Formal model · Self-preservation · Risk

1 Introduction

Deutsch (1958) identified numerous ways that make a human trust another human, yet models used by computational trust focus entirely on only one. This leaves a large area of human trust without proper computational representation. This paper proposes the computational model of a decision to trust that incorporates at least some additional ways of trusting. The model has been constructed with a view to explain the existence of trust in situations devoid of choice, but it is applicable to trust in modern technology as well.

What computational trust tends to consider is the situation of choice, where trustors can exercise their free will to trust one of several trustees, usually up to the level defined by their trustworthiness. This paper brings into the domain of computational trust those cases where trust seems to defy the rationality of a choice, specifically where it departs from simply reciprocating trustworthiness. These are the cases where trust exists despite having no choice and where trust does not exist despite the existence of valid choices. The

proposed model is simple, extends known models by being compatible with them, and addresses important cases.

1.1 Computational trust

Computational trust (Marsh 1994) is the research domain that models human trust in an algorithmic (thus computational) way. Its models (as well as the model presented here) approximate social and individual behaviour related to trust, in a view to support and replicate those behaviours with an aid of technology. For that end, it is important for computational trust to cover most of the situations where trust exists.

Research in computational trust concentrates on situations where the trustor (the one that trusts) has a choice, usually a choice to trust one of several readily available trustees. Such situations are common across the Internet and across the economy in general, where there may be even an over-supply of aspiring trustees such as Internet shops, opinion-making 'influencers' in social media or suppliers of information. Various algorithms, specifically reputation-based ones (e.g. Wierzbicki 2010), that aids the decision to trust, were studied and applied in this area.

This leaves situations with no or little choice under-represented throughout research in computational trust. Some researchers (Lewicki and Bunker 1995) even claim that trust always requires a choice, so that if there is no choice, one

✉ Piotr Cofta
piotr.cofta@utp.edu.pl

[1] University of Science and Technology (UTP), Bydgoszcz, Poland

cannot rightfully speak about trust at all. In the absence of other choices, there is at least a choice between trusting and distrusting (Luhmann 1979). Yet another (Castelfranchi and Falcone 2000) claim is that trust is a choice in itself as there is always an option of not trusting. Others (Cofta 2007) indicate that in situations of no choice, people may not trust but masquerade it with a trust-like behaviour.

It is, however, hard to ignore that there are several situations throughout our lives where there is no apparent choice yet the trustor acts as if there is an actual trust, and in retrospect describes such situations using the vocabulary of trust. There are also situations where the trustor withdraws from trusting despite having a choice, and even in situations where some of the trustees seem to be sufficiently trustworthy. Finally, there are situations where trust is placed in unreal or abstract objects despite no evidence of their trustworthiness.

1.2 Key proposition

This paper explores the concept that in all those cases listed in the previous section there is a force at work that has been widely ignored: self-preservation. Specifically, the decision to trust reveals the relationship between choice, benefits of trusting and self-preservation. Thus, the trustor trusts mostly because it improves its chances of survival, even though sometimes it may not be entirely comfortable with it. Decision to trust, even though sometimes perceived as irrational, is in fact rational. Therefore, this paper concentrates on rationally explainable decision to trust, i.e., on the model of process when the trustor has to resolve whether to engage in the further relationship with a trustee, and which one.

This paper accepts that the rationality of the decision is not contingent on such decision being logically processed or explained. That is, the emotionally driven decision can be as rational as the one driven by the logic. Every decision that supports the self-preservation can be considered rational (Karni and Schmeidler 1986), thus potentially algorithmically describable. The model proposed here follows the concept of utility maximalisation with self-preservation defining the von Neumann–Mongerstern utility function (von Neumann and Morgenstern 1953).

As already stated, the objective of this paper is to bring into the fold of computational trust those cases where trust seems to defy the simple rationality of a comfortable choice. As this is the propositional and exploratory paper, the author does not claim that this is the only valid approach, but rather would like to start and facilitate the discussion about the subject. Further, it is not within the remit of this paper to create an elaborate (hence complicated and detailed) model of trust. It is rather to propose the model that is simple, does not contradict known models and that addresses the majority of cases.

1.3 Contribution

This paper belongs to a stream of research that is concerned with the modelling of human decisions to trust, to allow for the algorithmic simulation of it. The main contribution of this paper is a proposition of the new model of decisions to trust. Main benefits of this proposition are as follows.

- The paper offers a simple model that links decision to trust with self-preservation through risk minimisation. This is a novel approach that builds on existing concepts and significantly extends the vocabulary and instrumentation of computational trust.
- The model is based on the theory of social systems, and is likely the first one that formalises the decision to trust between systems. As such, it can be hopefully applicable to the wide range of social relationships, including organisational trust, trust between people as well as trust on the Web.
- The model allows to explain phenomena that were largely ignored or ill-explained by existing research: trust in situations of no choice, trust in monopolies and similar. As those situations are in fact quite common in modern life, it hopefully makes computational trust more relevant to everyday experience. This paper discusses several motivating cases and then demonstrates how the model can explain the existence (or the non-existence) of trust in such cases.
- The model uses the single metric of complexity and a single function of risk to describe the process leading to the decision to trust. This makes the experimental verification of the model much easier, compared to existing models. Further, the mathematical formulation of the model allows for its further analysis and refinement.
- The model can be particularly applicable to relationships between people and technology, specifically to trust in technology and through technology. Modern technologies (such as social networks, blockchain or the phenomenon of fake news) tend to redefine our perception of trust and trustworthiness. The model can be used not only to explain some of the empirical phenomena, but also can guide the design of technology that facilitates trust and its development.

For clarification, it is the author's understanding that the computational trust should resemble the actual trust. That is, formulae provided by computational trust should provide solutions that are close enough (yet not exact in every detail) to those known from psychological or sociological research. To this end, the paper contains an extensive review of various strains of research that are relevant to this area.

1.4 Structure

This paper is structured as follows. Section 2 elaborates on thesis discussed in this paper. The next section discusses a set of motivating cases that will be used later to demonstrate the applicability of the model. Section 4 contains the review of the literature specific to the subject. Section 5 is an introduction to social constructivism focusing on aspects relevant to this paper. As the proposed model uses both risk and trust, Sect. 6 discusses the relationship between those constructs and their use in modelling. The next section discusses whether the construct of trust analysed in this paper is coherent with popularly used concepts about trust. Section 8 contains the formalisation of the model and demonstrates its applicability to motivating cases. The final section contains a concluding discussion.

2 Thesis and methodology

2.1 Thesis

The main thesis of this paper can be summarised as follows.

T1. There are always trustees to choose from, even if a trustor is not aware of them

The question of no choice due to the lack of trustees is not what it may look like. A trustor always has a choice, i.e., there are always some trustees that can be trusted, even though the existence of those trustees may not be obvious to the trustor while their suitability can be questionable. This thesis will be demonstrated by the reference to the constructivist approach to social systems.

T2. The decision-making process can be rationally described

Trust is often attributed to emotions, hormones or intuition, which makes it defy rational explanation. This is to certain extent correct, as trust can be easier post-rationalised than explained prior to the decision to trust. The model assumes that the mechanism used by the system can be described in a rational and simple way, with the potential to be for a formal and algorithmic representation. It is the same mechanism that drives decisions when there is a wide selection of trustees and when there is a very limited one, when the trustor is at the brink of self-destruction and when not. This will be demonstrated by introducing the formal model of trust-based decision-making.

T3. decision to trust a trustee is driven by a single driver

There is only one driver to trust: the threat to self-preservation, perceived either as an immediate threat or as a long-term improved potential to maintain self-preservation, as suggested by the theory of social systems. This will be demonstrated by a single model that operates on the notion of complexity and its reduction through trust.

T4. The model explains phenomena of trust that is defying a choice

Trust that defies a simple logic of a choice (that is, trust with no choice as well as no trust with a choice) is also covered by the formalisation. The discussion will be provided to demonstrate how the formalisation applies to such cases and how certain phenomena of trusting can be explained by the formalisation.

2.2 Methodology

The paper is structured as a theoretical research aiming at developing a new hypothesis. It follows a pattern of four steps, mapped into the structure of the paper. These are:

1. Identification of a problem. It starts in introduction (Sect. 1) and continues through motivating cases (Sect. 3) that are inspired by thesis (Sect. 2). Motivating cases form a foundation for the verification of a model.
2. Critical analysis of existing solutions. These include literature review (Sect. 4) and continues through Sects. 5 (about constructivism) and 6 (the review of relationship between trust and risk).
3. Synthesis of a proposed hypothesis. Initial formulation of the proposition is provided through thesis in Sect. 2, to familiarise the reader with the idea. They are then developed through the discussion about the nature of trust (Sect. 7), into a model described in Sect. 8.
4. Verification against defined cases. The verification is done in Sect. 9, against motivating cases (thus against thesis) and then against a real-world example. For clarification, the experimentation to verify the hypothesis is not in scope.

3 Motivating cases

Those motivating cases listed below illustrate a spectrum of situations that range from trusting with a choice to trusting with a limited choice or with no choice at all.

1. Nominal case. Trust with several choices.

A person resolved to book some hotel rooms over the Internet. She investigated several hotels and eventually settled for the one that had a slightly better reputation even though it was not the cheapest one. The reputation was warranted as the hotel turned out to be a good one.

2. Withholding trust as others are not trustworthy enough.

A pensioner resolved to let the professional investor take care of his lifetime savings, as making his own investment decisions became increasingly hard. He visited several investors, only to resolve that he does not trust a single one. He still keeps managing his funds all by himself.

3. Trusting more those who make a more compelling offer

There are two competing services providers on the market: one that promises only what it can deliver and the other one that promises to take care of everything. While the promises of the second provider may be somehow unrealistic, and the small print is particularly convoluted, it steadily gains the market share.

4. Trusting with limited choice.

A hospital patient has just found out that there is only one doctor who can perform a surgery, hence there is no choice of operators or hospitals. Pressed for time, despite the lack of choice, the patient apparently trusts the doctor, possibly only on the basis of doctor's impeccable bedside manners. For the sake of this case, the surgery went well and the patient feels that his trust was fully warranted.

5. Trusting non-trustworthy ones.

Students are aware that their favourite social media site manipulates the content, compromise their privacy and experiments with influencing their political opinions. Yet they still participate and trust that the site will eventually improve on its behaviour, because they feel that without trusting the site, they will be eliminated from the social life as they know it.

6. Trusting under duress.

A non-democratic government seized power in a country. For the lack of working political opposition, as well as for closed borders, this situation provides no choice to citizens. Some resolved to trust the new government, against their beliefs. They justify it by stating that at least the government upholds the letter of the law, no matter how unfair the law is.

7. Choice and trusting.

It has been reported that patients who have a choice of their primary care physicians generally express more trust that those who have physicians assigned with no choice (Kao et al. 2001). Still, if possible, they stick to the most trustworthy one. Thus, the existence of a choice influences the level of trust even if trustworthiness is not affected.

4 Literature review

The question of trust without a choice, as well as the impact of self-preservation on trust, falls between sociology, social psychology and psychology, with implications on (among others) the design of information systems. This literature review has been conducted mostly from sociological perspective, but it touches on some psychological aspects as well. Note that this paper contains also smaller literature reviews concerning more specific problems.

Psychologically inspired models of trust tend to focus on the decision to trust. McKnight and Chervany (2001) introduced the model that identifies four components that influence our decisions to trust: disposition to trust, institution-based trust, trusting beliefs and trusting intentions. Collectively, they form a transactional trust that influences the decision whether to engage in a transaction. The model covers only the transactional relationship between a trustor and one trustee. The choice available to the trustor is to trust a trustee or not.

Tan and Thoen (2000) introduce model that focuses on transactional trust, where the trustor engages in a transaction after considering the potential gain and risk, taking into account trust in a trustee as well as trust in control mechanisms that keeps this trustee in check. The model assumes that there is certain threshold of trust that is required to engage in transaction, and if there are no reasons to trust, the trustor should not engage. While the model (same as the previous one) does assume a choice (at least a choice between engaging and not engaging), it introduces the notion of risk and payoff into trust-based decisions.

Regarding the availability of trustees, psychology knows an interesting phenomenon of an 'imaginary friend' (Klausen and Passman 2007) that children engage with instead of real playmates, possibly to bring some comfort. The phenomenon usually passes with growing maturity, but it is an important observation that human mind can (and often does) freely create personalities that feel real to someone.

The relationship between self-preservation and risk can be described using the concept of rational utility maximisation, where the utility function maximises short-term

survival, possibly at the expense of longer-term opportunities (Karni and Schmeidler 1986). Some aspects of this approach are present in this paper while discussing the relationship between trust, risk and self-preservation.

For sociology, choice seldom appears in research papers, even as a side-line interest, as if the authors assume that the society and its components always have a choice. However, specific areas such as trust in monopolistic organisations (or governments) can at least indicate the general approach to situations of no choice. Still, it is often necessary to refer to the overall concept presented by the authors that by the particular passage in their texts, and only infer what is their view on trust without a choice.

Luhmann addressed trust as a social phenomenon twice (1979, 2005), from slightly different theoretical perspectives. This paper accepts that both views on trust complement each other. Trust is a social mechanism that minimises the complexity of decision-making by catering for uncertainty. It is also a way of a social system to off-load some of the complexity it deals with. Trust emerges through the phenomenon of dual contingency where both systems learn to rely on each other. Luhmann does not directly address the problem of a choice and trust, possibly because in the domain of social systems, one system always has some choice (i.e., there are always other systems).

Still, there are some observations made by Luhmann that are useful for this paper. First, he observes that trust is a necessity of modern life, i.e., that "a complete absence of trust would prevent him [a person] from getting up in the morning" (Luhmann 1979, p 4). This indicates that trust is often the default option, no matter whether there is a choice or not, as otherwise we end up in a situation of "chaos and paralysing fear" (ibid. p. 4).

Second, trust is concerned with contingencies: trustor is aware that things can go wrong and granted trust, in retrospect, many not have been warranted. Thus "people, just like social systems, are more willing to trust if they possess inner security…" (ibid., p 78). This indicates that the situation of a trustor is important for trust to be granted.

Next, trustor "is never at a loss for reasons and is quite capable of giving an account of why he shows trust" (ibid. p 26). That is, the rationalisation of a decision to trust is a volatile concept, as trust can be always post-rationalised, specifically if things go wrong.

Finally, it is worth reinforcing the original notion of trust as complexity reduction: "Trust is rational in regard to the function of increasing the potential of a system for complexity" (ibid. p. 88). This may imply that any action that allows the system to deal with increased complexity warrants trust. This observation is one of the foundations of this paper.

Hardin (2002) strongly binds trust with trustworthiness, and argues that the most common form of trust is the one that asks for the encapsulation of interests of a trustor and its trustee. Rational trustor should determine whether trustor's best interest is encapsulated in trustee's best interest and vice versa. The question of trusting without a choice does not seem to be present in his works, which is not surprising, as the concept of encapsulation of interest presupposes both a trustor and a trustee exercising at least some elements of free choice.

In this context even more interesting is Hardin's observation on trust in government, as citizens cannot switch from one government to another at will, making the situation closer to trust with limited choice (assuming that emigration and re-election are viable but expensive). He notes that in such cases interpersonal trust, as well as the concept of the encapsulation of interest may not apply.

Instead, he proposes the concept of one-way quasi trust, which is the combination of trust in capacity but not necessary in intentions of the government and the rational decision to reduce trust to the considerations of risk. The necessity of trusting the government is then down-played, so that for as long as the government is not actively distrusted, the relationship between citizens and its government may continue.

Möllering (2006) provides a wide overview of various approaches to trust and trustworthiness. While the structure of his works does not allow to deduct author's own thoughts on trust and choice (apart from the observation that trust differs from the rational choice theory, yet it is still rational), there are several relevant observations within the text.

Becker's approach (2005, after Möllering 2006) is discussed in relation to pragmatic approach to trust, where the person willingly trusts in what he believes is true, where 'true' has a pragmatic meaning of being useful, giving expectations, and enabling actions. While the author relates this approach mostly to the way the trust process can be started (after all, someone has to trust first without any evidence), it can be generalised on situations where there is nobody worth a trust, yet trust is necessary.

McKneally et al. (2004, after Möllering 2006) analyses the behaviour of patients who underwent elective surgeries. Surgery always requires trust, specifically elective ones where the patient has certain control over his (or her) destiny. Still, patients developed several psychological mechanisms to deal with their doubts even if evidence or choice was sometimes scarce.

Giddens (1991) discusses ontological security and basic trust that relate to the perception of the continuity and the stability of the world. These are the basic foundations of trust as well as of self-identity. They are formed in the early childhood, and the violation of those may lead to the significant damage to self. Notions of 'disintegration' of self or the 'inability to reconstruct itself' used throughout this paper closely relate to those concepts used in his paper.

Further, Giddens (1990) discusses the role of abstract systems in modern society and our trust in such systems. He claims that abstract systems (such as the law etc., but also any expert system where the knowledge is hidden from the layman) disintermediate direct contacts between people while making some people 'access points' to such systems. Despite their near-monopolistic position they are the focal point of our trust relationships. Within the scope of this paper, there is a similarity between those observations and the concepts of social systems (specifically organisations and function systems) from Luhmann (1995).

Kasten (2018) reports that trustworthy behaviour is indicated by three kinds of behaviour: shared social identity, socio-emotional needs of the trustor and compliance with social norms and obligations to trust. This paper relates to those concepts. Thus, shared social identity relates closely to concepts such as the encapsulation of interest and interpenetration that is leading to shared meanings. Satisfaction of socio-emotional needs is close to the concept of a penalty for not having a choice. Finally, compliance with social norms relates to situations where, in the absence of first-hand evidence, trust is drawn from abstract systems.
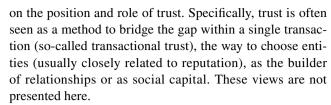
Williamson (1993) claims that several situations that are casually attributed to trust can be explained by calculative approach. This claim, still debated in its accuracy, is specifically important if we consider that commercial organisations are social systems and that such organisations may value calculative risk-taking over non-calculative trust. Thus, one can debate trust and risk using the similar framework.

Ward and Smith (2003), while discussing trust in business relationships make an interesting observation about trust and choice by differentiating between types of trust. Thus, authentic trust (i.e., trust willingly developed between entities) is always a matter of choice (of trusting or not) and has to be given freely. Network trust (trust within groups) limits a freedom of choice, but not entirely. Authority trust used to eliminate an element of choice, but more recently authorities can be picked at will. Finally, commodity trust (casual trust in business relationships) is seen as a 'take it or leave it' case, where no growth of trust can take place.

5 Constructivism and trust

This paper is based on social constructivism, specifically on works of Luhmann (1995) and the theory of social systems. It assumes that trust is a way used by social systems to deal with their growing complexity. This section will provide an introduction to constructivism and will also address the question whether there is always a choice and there is always someone (or something) to trust.

Constructivist's approach remains close to the role of trust between organisations. There may be alternative views on the position and role of trust. Specifically, trust is often seen as a method to bridge the gap within a single transaction (so-called transactional trust), the way to choose entities (usually closely related to reputation), as the builder of relationships or as social capital. These views are not presented here.

Constructivists distinguish between two ways trust is used: as a way of building mutual dependencies between systems and to minimise complexity of a system by moving (off-loading) such complexity to another system. In practice, both forms often reinforce each other. Thus, the working relationship makes one system more likely to move its complexity to another system while positive experience with such move reinforces the acceptance of such relationship.

5.1 Social systems

Social systems theory assumes that the social reality is constructed from (and of) communications that are grouped into social systems. Each system operates within its environment that contains all other systems. People contribute to systems by processing communications, but systems exist among people. Systems continuously accept and integrate into themselves new communications that help them self-recreate their structures through the process of autopoiesis.

System operates by responding to new communications that arrive from the environment of the system by creating some communications, and all those communications become parts of the system. Thus, every communication generates more communications from and within the system, making other systems generate new communications.

Systems cannot prevent the environment from generating new communications and cannot ignore them. Therefore, the number of communications that the system contains (i.e., its complexity) grows all the time. Because communications generated by the system should relate to all communications that constitute the system, the larger the system is, the slower the response will be. Therefore, every system must engage in a form of complexity reduction or face the threat of extinction by irrelevance—it will stop being able to respond in a timely manner so it will be no longer relevant to its environment. While systems should not be treated in an anthropomorphic manner, the evolution of systems left only those that somehow took care of their self-preservation.

There are two ways of managing growing complexity and assuring self-preservation: constructing meanings and interpenetration, the latter one being associated with trust.

Meaning is a selected cluster of communications within the system that guides the responses instead of all the communications the system consists of. Whenever the new communication arrives, the system should analyse all its communications and figure out the response. As this process can entail too much complexity (and slow down the system

beyond relevance), the system constructs the meaning as a synthesis of all its communications, in such way that the meaning can be examined rapidly whenever the communication arrives. Drawing an analogy to information systems, the meaning can be considered a simple internal 'state' of the system that guides the response of the system. This avenue is always available to the system provided that it is able to produce (and modify) meanings in response to changes in their environments.

Interpenetration is a method employed by the system where the system 'exports' its communications and 'imports' ready-made meanings from another system. This mechanism resembles—allowing for another crude analogy—outsourcing data processing to subcontractors and getting from them analytical results. This process requires what may be colloquially called a trusted subcontractor—a trustee who will not abuse the privileged position and will act in the best interest of a trustor. Otherwise, the system will end up with incorrect meanings and without the ability to correct them, making the system increasingly irrelevant. This is the core mechanism discussed in relationship to trust and trustworthiness.

It is the main tenet of this paper that it is only rational for the system to weight the risk of extinction against the risk of trusting (thus interpenetrating) systems that are not as trustworthy as it would like them to be, specifically if choices are limited or non-existent. The decision is the outcome of the weighting of the threat of extinction, inability to construct meanings and availability of trustworthy systems. While trust comes to a decision that is frequently made under duress, under threat and in situations of scarcity, it is trust nonetheless.

5.2 Monopolistic systems

Monopolistic systems play a special role in this paper, as they are very popular, both in a social domain (e.g. the legal system, the government, the army, etc.) and in the technical domain (e.g. monopolistic utilities provider, monopolistic Web service, etc.). This section concentrates on demonstrating that existing theories of trust do not preclude trusting monopolies.

Trust, according to Luhmann (1995), emerges from double contingency, i.e. from situations where both a trustor and a trustee understand what the other system wants and provide what it wants, even though they both can do otherwise. In a way, they attune their meanings to each other's needs in a process that is gradual and voluntary. This view has been also conceptualised as the 'trustor's debt' by Coleman (1982). This mutuality of trust reverberates also, e.g. in Hardin's (2002) encapsulation of interest. Thus, trust is always 'between' and never unidirectional, which is not how it usually works with monopolists.

When it comes to monopolies and trust, Luhmann indicates that ".. there are functionally equivalent strategies for security and situations almost without freedom of choice, for example, in the domain of law and organization.." (ibid. p 129). Those strategies are based on trusting systems that can control the trustee, instead of trusting directly the trustee. This relates to the concept of using abstract system (Giddens 1991) as a vehicle to develop trust, where trustor, by trusting such abstract systems, can utilise their controlling power as well as can rely in symbols that such systems produced.

This, however, only deflects the problem instead of resolving it. Abstract systems that have the ability to control are often monopolistic as well, with the system of law and law enforcement being primary examples. Hence this approach replaces trusting the potentially unknown single trustee (e.g. Web provider) with trusting the known monopoly (e.g. international law enforcement). The only difference here is that the trustor is more likely to have an experience with such controlling systems (e.g. with the legal system) while it may not have any direct experience with the trustee (i.e. the Web provider).

Indeed, the problem of our ability to trust monopolists or, more generally, to trust when there is a significant disparity of power, is still puzzling researchers. The problem itself is not only relevant to technology-based monopolists, as it can be generalised on the level of archetypal trust in trust the world around us, as this world is the natural monopoly without much of the choice—as it is the only one available to us.

This paper suggests that monopolies can be trusted, but the psychology behind it differs from interpersonal trust. The initial observation comes from the area of politics, where trust in government (being a monopoly) is discussed,. Hardin (2002, p 151) says "One might still wish to say […] that a citizen can trust government, but the 'trust' in this case is almost certain to be different from the trust I might have in you." In discussing how different such trust may look like, he provides two possible directions.

The first one is to use abstract systems to control the monopolist, in a manner already discussed. Considering the problem of trusting abstract systems, he observes that (Hardin 1991) "My trust in 'the market' may be like my trust in the sun's rising tomorrow". This direction is further reinforced by Giddens (1990), who discusses the role of ontological security and basic trust related to our perception of the continuity and the stability of the world.

The second direction comes from Hardin's observation that in case of monopoly "… we should generally speak not of trust in government but only of confidence…" (2002, p 172), referring to Luhmann's (1988) distinction between familiarity, confidence and trust as different modes of asserting expectations.

Familiarity and confidence presuppose asymmetric relations between system and environment. Familiarity indicates that the system incorporated the part of the environment into itself so it knows what to expect. Confidence indicates

the system does not know the environment, and having no impact on the environment it only expects from the environment to keep behaving the way it does. For completeness, trust presupposes that the system and its environment have a mutual impact on each other.

Both lines of thought lead to similar conclusions: one can trust (or be confident, or feel secure) in a monopolist (including the world as a whole) by accepting the monopolist as it is, and by being content with observable regularity in the behaviour of the monopolist, even if this behaviour is not always beneficial. In this process, the persons self-preserve their psychological integrity, which is more important to them than the perception of a 'genuine trust' (Solomon and Flores 2003).

From this relatively long overview it is clear that the sociology and social psychology do not exclude the possibility of trust between small systems and large monopolies where choice is limited or non-existent.
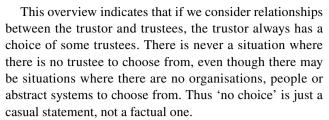
5.3 Availability of trustees

This section concentrates on demonstrating that existing theories explain how there can be always a trustee available.

The theory of social systems identified three main classes of systems: organisation, functions and interactions (Luhmann 1995). Of those, organisation systems mostly resemble what we may call 'visible entities'—i.e. distinguishable trustees that a trustor can relate to. Function systems (law, media, etc.) permeate the social reality and are maybe harder to relate to, yet their existence cannot be denied. Large organisations, specifically ones that embody function systems, called sometimes abstract systems (Giddens 1990) are the visible example of the organisational aspect of function systems. Interactions are systems that are both more pervasive and less visible, as they focus on efforts to stabilise meanings. Still, discussions about some meanings (specifically evocative ones such a 'justice' or 'truth') are both pervasive and easily identifiable in the social sphere.

Consequently, no social system is truly alone—there are always some social systems in its immediate environment. The same can be said about people and social systems—even on the remote island one is always a part of several social systems and can relate to other systems.

Further, from psychological perspective, there are entities that are created by people for their private or semi-private consumption, such as 'imaginary friends' (Klausen and Passman 2007) or fairy tale stories and characters, with some of them being promoted to the shared social sphere. If nothing else, they demonstrate that the trustor can have a relationship with trustees that not only do not exist in a physical sense, but also are non-existent according to the prevalent logic.

This overview indicates that if we consider relationships between the trustor and trustees, the trustor always has a choice of some trustees. There is never a situation where there is no trustee to choose from, even though there may be situations where there are no organisations, people or abstract systems to choose from. Thus 'no choice' is just a casual statement, not a factual one.

Therefore, it is possible to reformulate the question of no choice: the trustor always has a number of choices, even in the most solitary situations. For example, let us consider a situation where the trustor desires some service, and arranging this service by himself is too complex. In this situation, there may be a choice of providers willing to deliver, but failing that, the trustor may resort to trusting the government, neighbours, abstract concepts of fairness, etc. Failing that, the trustor may imagine a trustworthy trustee that will deliver what is desired in a 'make-believe' fashion. Maybe not all of those solutions will satisfy physical needs of the trustor, but all of them may contribute to self-preservation.

Additionally, the trustor always has a choice of not trusting anyone. Indeed, as discussed later, this choice may be illusory if not trusting means facing the threat of disintegration, but it is a choice nevertheless. Again, 'no choice' is in fact just a casual shortcut for saying that there are only undesired alternatives, not that there is none.

6 Risk and trust

This paper uses the construct of risk to explain the inner working of the system that results in trust (or the lack of it) other systems. The proposed model states that the system chooses the path of lesser risk, and if such path requires trusting, then it exhibits trust. The introduction of risk to explain trust requires additional explanation, provided in this section.

6.1 The relationship between risk and trust

Despite the fact that the relationship between risk and trust has been a subject of several studies the outcome is far from being clear. Mayer et al. (1995) state that it is even unclear whether risk is an antecedent to trust, or is an outcome of trust. That is, whether trust requires a situation of risk that it helps overcome, or whether trusting creates a situation of risk. Regardless, it is the action of trusting, not the intention to do it that creates the risk (Deutsch 1958).

The primary difference between risk and trust is that trust can be approximated as a subjective probability of a success due to the quality of identified actors (Gambetta 2000) while risk is an objectified probability of failure due to identified actions (Nickel and Vaesen 2012). That is, trust is the

subjective state of mind, the decision and the action of the trustor who believes that by trusting an identifiable trustee, it will be more likely to reach some desired state. As trust is subjective, the mental process of trusting is potentially complicated (Castelfranchi and Falcone 2000).

Contrasting, (Nickel and Vaesen 2012) risk is a perception of a possibility of future harm coming from some actions, supported by quasi-objective evidence. Risk is expected to be estimated through statistical processes that draw from historical data about similar actions from the past. This leads to a common practice to express and process risk through probabilistic formulas.

Jøsang and Presti (2004) analysed the relationship between risk and trust in the context of decision-making. They confirm that both risks and trust can be used to make decisions in the uncertain environment. Both risk and trust should affect the actions of the rational agent and can be incorporated into a model of decision to trust.

Luhmann (2005) observes that risk is the primary modern way of thinking about the future. That is, the future course of actions may cause risks, and the rational actor (be it a person or an organisation) should consider all the possible risks, and their possible configurations. This may include the risk of not acting or the risk of risking. This approach is restrictive, as it removes the positive drive provided by thinking in terms of trust.

Bohnet and Zeckerhauser (2004) studied the relationship between risk, trust and betrayal, using the game-theoretic approach. They observed that the assumption that the willingness to trust is being closely related to the willingness to take risks, has not been empirically validated. Their experiment demonstrated that betrayal means very little to those who take a risk-based approach while it represents significant loss to those who took a trust-based approach. That is, risk incorporates the chances of betrayal into expectations while trust does not.

Galli and Nardin (2003) explored the role of trust as a risk moderator, where trust allows people to pursue courses of actions that are otherwise considered too risky. It is interesting that the level of perceived risk correlated with the level of complexity of the task. They found out that the role of trust is not that important at low-level risk, but trust becomes important when tasks are risky.

Differing from uncertainty, risk tends to be defined as a situation where there are known chances of negative course of actions due to actions or inactions of specific entities. That is, while it is not known whether those entities will or will not behave in an unfavourable way, at least the probability distribution (or other measure) of negative outcome is known. Risk is thus distinguished from uncertainty where such probability distribution is not known and from danger where entities are not known (Luhmann 2005).

The knowledge of probability distribution leads to the quantification of risk in a form of expected loss value (ELV) (Chapple et al. 2018), which combines the quantification of loss with the quantification of probability. In some areas the ELV is considered to be the value of risk or the risk itself.

The approach taken in this paper is close to the concept of trust as a rational risk-taking, approached mostly from game-theoretic perspective (Nickel and Vaesen 2012). The assumption is that the rational agent should follow the utility maximisation by adjusting its behaviour to the known distribution of risk.

6.2 The use of risk in models of trust

Risk is incorporated into known trust models in several ways, and this paper follows this line of thoughts. The objective of this section is to demonstrate how risk can be integrated into the reasoning about trust.

Mayer, Davis and Schoorman (1995) introduced the integrative model of organisational trust that combines the perceived trustworthiness of a trustee, trustor's propensity to trust and perceived risk. The model is not computational, only descriptive and qualitative. They state that "risk is an essential component of a model of trust" and consequently the notion of risk is used across the whole model.

Propensity to trust is a personality trait regarding the generalised willingness to trust others and sets the default (initial) level of trust in any relationship. This propensity is closely related with the trait of risk tolerance, i.e. the generalised willingness to engage in situations with higher risk (Sitkin and Pablo 1992). The key difference is that propensity to trust is stable across several situations while risk tolerance is more situation-specific.

The role of perceived risk in the model of trust is to reaffirm that trust is needed only in risky situations—if there is no risk, then trust is not necessary. This component models what is often called the 'transactional risk'—the unavoidable situation that the trustee, after all, may default. The trustor must understand that there is a risk and yet accept it due to trust.

Finally, the decision to trust (which is also the decision to take the associated risk) depends on the balance between the perceived trust and the perceived risk. If the level of trust exceeds the level of risk, the trustor will act to trust, if not—the trustee will not engage. Thus, the whole model is based on setting the threshold value that has to be achieved by trust.

McKnight and Chervany (2001) introduced one of interdisciplinary models of trust applicable to electronic commerce. The model comprises of four components: disposition to trust, institution-based trust, trusting beliefs and trusting intentions. The model itself is an exception in that it does not refer to risk, introducing it only marginally through the notion of the possibility of negative consequences in its last component.

However, their extended model (McKnight et al. 2003) introduces the notion of risk in several ways. It observes that there are in fact two dispositions: to trust and to distrust, driven by the level of risk: trust is associated with low-risk situations while distrust is associated with high-risk ones. The model therefore suggests that, before engaging the interaction, the trustor performs a risk analysis that affects the whole trust-related process.

The same extension of the model openly admits that it is the risk of an interaction that has to be met with trust to make the interaction happen, thus reaffirming the construct of a threshold. Specifically, while discussing the interaction through the web site, the model distinguishes between low-risk situations (such as a casual browsing) from high-risk situations (such as engaging in transaction), demonstrating that these are driven by different causations.

Tan and Thoen (2000) model decision to trust in electronic commerce, specifically addressing trust in e-commerce transactions. The model is built around a similar notion of a threshold, defined by risk and expected gain in transaction (effectively the risk-adjusted gain). Trust arrives from two components: direct trust in a trustee ('party trust') and trust in control mechanisms that contain the trustee ('control trust'). Transaction is likely to happen when the total trust exceeds the threshold defined by the risk.

The model directly refers to the notion of risk while defining the transitional threshold. If nothing else, it indicates that metrics of trust and risk are comparable, as otherwise it would not be possible for the threshold to operate at all. Note that this model introduces the duality of trust and control that itself deserves a discussion, as instruments of control often relate to risk-based measures. For an extended discussion see (Cofta 2007).

Reputation-based models of trust (Jøsang et al. 2007) tend to focus on converting subjective assessments of one's trustworthiness into quasi-objective reputation, in expectation that such reputation will be taken as an estimator of transactional trust. As models of trust-based decision-making they have two distinguishing features: they introduce choice and do away with threshold.

When reputation is concerned, the decision of a trustor is not limited to a single trustee, but rather to the choice among several available trustees. That is, the trustee can learn about reputation of several entities and then make a decision whose entity is worth its trust. Thus, reputation-based schemes are in fact decision-support tools, not decision-making models.

When it comes to threshold, the trustor is not facing the go/no-go decision characteristic to the threshold-based models. Instead, the trustor can decide what is the risk level that it accepts while dealing with certain level of reputation, and seek the trustee that operates at this level.

Thus, e.g. trustees of higher reputation may sell goods at higher prices (Resnick 2006).

6.3 The use of risk in the proposed model

The model presented in this paper deals with situations where a trustor can engage in some forms of relationship with one of several available trustees. This model uses the construct of risk to explain the inner working of the trustor that results in trust (or the lack of it). The model states that the trustor chooses the path of lesser risk, and if it requires trusting the trustee, then it exhibits trust.

Risk, as a construct, is multi-dimensional, with several colloquial usages but also with several strict domain-specific definitions. The defining characteristics of risk are the known probability of negative outcome (as different from uncertainty) (Borch 1967). However, it is not certain whether 'risk' refers to the state where such an outcome is possible, its probability or the overall probability-adjusted loss associated with such an outcome.

Risk, as used in this model, is defined as a probability of a negative outcome of a given action. The lack of the use of probability-adjusted value is justified as what is at risk is the existence of the trustor, that has an unlimited value to the trustor and that is constant across the whole model.

Within the model, risk is expressed through risk functions that provide the assessment of such probability for a given set of circumstances. Those functions encapsulate the parametrisation of the assessment of risk, as well as its potential subjective elements.

The model does not imply any particular relationship between trust and risk, as it does not mix trust and risk while explaining how the entity makes a choice. Trust and risk are only connected by the decision to trust, so that risk models the behaviour of the entity that leads to such a decision while trust models the outcome of this decision.

Risk has been chosen to model the operation of the entity mostly because it is a common way of thinking about the future (Luhmann 2005). While it is unlikely that there is a common way of describing how different entities come to the decision to trust, the use of risk makes it more likely. If nothing else, risk is therefore a convenient method of description.

Trust, in turn, is used by the entity because it considers its own future in a way that risk cannot. Trust is necessary for the entity to survive, as it has to engage in meaningful relationships with other entities.

7 Is this really trust

This paper concentrates on externally visible actions of trusting: a trustor will subject itself to vagaries of a trustee, what may be considered to be a non-genuine trust. This section

briefly reviews what is genuine trust and how it differs from calculative thinking and a resignation. It demonstrates that while trust is not always genuine, it is trust nonetheless.

Literature tends to distinguish between the internal state of trusting and the external behaviour that is identical with trusting (Cofta 2007). Solomon and Flores (2003) specifically distinguish between the authentic trust where mutual trust is freely and willingly granted and situation where trust is unilateral, not though through, forced under duress or otherwise non-authentic.

In a similar manner, Deutsch (1958) makes a distinction between a genuine trust that reciprocates trustworthiness and several others sources of trust such as despair, conformity, innocence, impulsiveness, virtue, masochism, faith, gambling or confidence. However, he does not condemn those non-genuine forms of trust, indicating that they belong to the spectrum of human behaviour.

An interesting alternative in thinking about various forms and sources of trust is provided by Vanderhaegen (2017) in a form of dissonance (Festinger 1957) and its control. Dissonance is a state where the system finds itself in a situation where communications it receives from its environment are in conflict with the meaning it already developed. Long term, dissonance has a negative impact on the system. The system can provisionally tolerate the dissonance, even if it leads to the development of two or more competing meanings. It can also reduce the dissonance by altering the meaning or by altering the environment or by rationalising the discrepancy.

Framing the problem of genuine and non-genuine trust as a problem of dissonance, one can see that the 'genuine' trust that reciprocates trustworthiness can be conceptualised as the stable state where no dissonance occurs between internal considerations and external actions. Contrary, for other sources of trust (the non-genuine trust) there is a risk of dissonance, as internal meaning and external actions diverge. Such a dissonance can be identified and to certain extent controlled for the benefit of self-preservation.

Another approach may be to view the non-genuine trust as a calculative action. The question of calculativeness and trust has a long history (Möllering 2014) and it is unlikely that this paper will close this discussion. This paper takes a stance similar to Williamson's (1993), that trust is a cost-effective safeguard, stressing that internally calculative operation leads to accepted external dependency built on trust.

Along similar lines, Harvey (2013) states that trustor can be calculative in decision to trust, yet the trust is genuine as it bears the possibility of betrayal. Thus, what defines the genuine trust is not the internal state of a trustor, but the fact that the trustor, once deciding to trust, can be hurt by the trustee. Which is the case for this model.

Then there is a question of resignation: whether one with limited choices really trusts or simply resigns to it. The argument here is that this is not the valid formulation of this question. Trustor always has an option of not trusting, if he is willing to bear consequences of it. However, if the trustor calculated that those consequences are too severe, then trust is genuine in the sense of Harvey (2013).

From psychological perspective, people can always convince themselves about validity of their trust (or distrust) ante-factum and post-factum, through the tendency of a choice-supportive bias (e.g. Mather et al. 2000). Evidence can be re-interpreted and memory can be distorted to justify a choice. As an example, Butler et al. (2012) demonstrated how incentives alter reported beliefs, once expectation of beliefs has been established.

Definitions of trust do not provide a definitive response to a question of genuine trust. Generally speaking, the meaning of trust, as all meanings, is managed by social interaction. That is, definitions of trust (and perceptions of what constitute the genuine trust) may vary among disciplines and may change in time.

Definitions of trust (Mayer et al. 1995) link internal and external operation of the system, e.g. they require actual vulnerable dependency (externally visible) being taken willingly (internally). On the other hand, social systems theory (Luhmann 1995) is not concerned with internal states at all: if the system makes itself dependent then it trusts the other. 'Trust' is therefore a label that describes certain behaviour of a system. This is the approach taken by this paper.

The author does not claim that this is the one and true view on trust. Contrary, social systems theory clearly indicates that it is not the case. However, the author claims that such a view can provide a valuable insight on the observed phenomenon of trust in situations with no or limited choice.

8 Proposed formalisation and model

This section deals with the proposed formalisation of a decision to trust, in a form of a model, to demonstrate that decisions to trust and a choice of a the trustee can be expressed as a relatively straightforward process that is controlled only by a handful of variables.

The model introduced in this paper, while borrowing from established models, differs from them, as follows:

- Choice, not a binary decision. This model is closer to reputation-based models in discussing the choice that the trustor has, beyond the simple go/no–go decision.
- Decision model, not decision-support tool. This model attempts to explain the process of making the decision to trust, unlike models that leave the final decision to a trustor, being a decision-support tool.
- Best option, not a threshold. The model does not set any threshold to trust (or to risk), but assumes that the trustor will take the best option available.

- Self-preservation, not trustworthiness. Trustworthiness of trustees, even though included in the model is secondary to the needs of self-preservation experienced by a trustee. Trust is a tool for such self-preservation, not a vehicle to reaffirm trustworthiness.
- More situational, less contextual. The model highlights that the decision to trust depends on the situation of a trustee more than on the context of a transaction. Therefore, sometimes contextually 'bad' decisions can be situationally appropriate.

8.1 The model and its formalisation

The main proposition of the formalisation is that a trustor, as any other system, has to manage the risk to its self-preservation and survival, as it is under the continuous threat emerging from its growing complexity. Such a complexity emerges from the fact, that the system cannot stop accepting communications from its environment yet it has to respond to such communications in a timely manner. The risk is genuine as the trustor that stops being a real-time processor of communications and stops responding to communications from the environment gradually disintegrates. In the extreme case, the trustor may even disappear.

As it has been already mentioned, the system can respond to the growing number of communications by building meanings that will serve as shortcuts to guide its responses. This, however, requires the system to spend some of its processing on building such meaning. Further, dissonance (Festinger 1957; Vanderhaegen 2017) may lead to situations where systems have several incompatible meanings, so that the complexity is effectively not reduced as significantly is it could be.

As the social systems theory indicates (Luhmann 1995), the trustor can also limit its own complexity by exporting some of this complexity to other systems. As the trustor becomes vulnerable to vagaries of the trustee it exported to, it seeks entities that are willing and trustworthy and expresses its trust in them by the act of exporting complexity. However, exporting is not cost-free to the trustor either. It has to manage a relationship with the trustee, so that the trustor also has to import some of the complexity from whatever trustee it chooses.

Note that the notion of 'complexity' used in this model does not refer directly to the number of communications that constitute the system or that flow into the system. It is rather the measure of 'computational load' or 'computational complexity' of the task of staying current. Thus, there may be systems of higher processing capability that can take a higher load, but taking a load decreases the amount of available processing power, thus decreasing the ability to take further load.

8.2 Entities

The model discusses 'entities' that describe social systems as well as the people, as the theory explains the behaviour of both. Let n be a number of entities that the trustor can identify within its environment, and possibly already exported some of its complexity to. We denote the set of entities used by the model as $E = \{e_1, \ldots, e_n\}$, where e_i represents a single entity.

Note that for the constructivist there are always several entities to choose from, even though it may not be apparent, so it is always true that $|E| > 0$. There is never the situation when there are no entities to choose from, even though there may be situations where there are no entities of a particular quality.

8.3 The structure

Fig. 1 illustrates the overall structure of the model. As the model is concerned with the decision to trust that the trustor can make, it assumes (top row) that the trustor considers exporting certain complexity and that the trustor is in the specific situation when it comes to its own complexity.
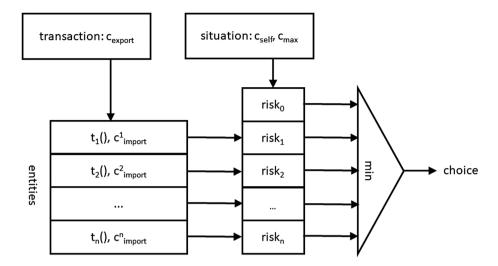
The trustor explores several potential trustees (left column), each one characterised by its trustworthiness and by the burden of complexity associated with relationship. The trustor assesses how the potential transaction may affect the trustor's own risk to its existence (central column), against taking no action at all (the first cell of the central column). The final choice is driven by the minimisation of risk (right-hand side).

8.4 Formalisation

The model assumes that the trustor has some current complexity c_{self} and certain maximum complexity that it can handle, c_{max}. As trustor's complexity grows, it faces the increased risk of failure (eventually leading to self-destruction). This risk is defined as the probability of failure to respond in time, and the value of risk is rapidly reaching one (i.e. the certainty of failure) as the c_{self} approaches c_{max}. We will denote this risk by a function $r_{self}(c_{self})$ which is likely to be parametrised by c_{max}. While the shape of this risk function is specific to the trustor, the model assumes that it monotonously grows, reaching one when $c_{self} = c_{max}$. Further, it assumes that the growth is non-linear: the function grows slower for small values of c_{self}, then accelerating its rate of growth as c_{self} is reaching c_{max}. This reflects the observation that the system should be truly concerned with its complexity only when it has little complexity to spare.

Assuming that the entity (the trustor) cannot anymore handle the complexity through the internal construction of meanings, it has to export some of its complexity to other

Fig. 1 The model of the decision to trust

entities. We will denote the amount of complexity that the trustor is willing to export as c_{export}.

The success of exporting depends on the trustworthiness of the trustee, a subjective probability that the trustee can handle the complexity to the satisfaction of the trustor. We will describe trustworthiness of a trustee e_i as a function $t_i()$, and it relates to the concept of 'transactional trust' (Tan and Thoen 2000). Being subjective, the shape of function $t_i()$ is specific to the relationship between the trustor and the particular trustee. Specifically, the function is likely to be parameterised by the current state of relationship, observable qualities of the trustee, amount of complexity exported, etc.

Note that the total volume of complexity that has been exported over time is not bound by the complexity of the trustor, it is only the c_{export} that cannot be larger than c_{max}. In fact, it is likely that the trustor, being a member of a modern society (Giddens 1990), has already exported much more complexity than it can handle by itself through several operations like this.

Exporting complexity does not mean that the trustor disposes of all the exported complexity, as even maintaining the relationship with another entity takes up some complexity. The same can be said about possible meanings that the other entity returns in exchange for the exported complexity. The amount of such complexity varies between entities and is denoted here as c_{import}^i. Thus, the whole 'contract' that the entity offers can be formulated in terms of export and import of complexity.

The operation of exporting the complexity changes the position of the trustor when it comes to the risk to its self-preservation. Instead of handling c_{self}, it ends up handling $c_{\mathrm{self}} - c_{\mathrm{export}} + c_{\mathrm{import}}^i$. However, this is the case only if the trustee turns out to be trustworthy, i.e. with the probability defined by $t_i()$. Otherwise, with the probability defined as $1 - t_i()$, the trustor not only has to handle its original complexity c_{self}, but it still has to deal with the penalty of

being betrayed by the trustee. The model assumes that this penalty will be of a range defined by c_{import}^i. That is, the trustor must continue some form of the relationship with its former trustee.

If the trustor executes a simple strategy of risk minimisation, the trustor faces a simple choice of selecting one of entities (or none at all), where the overall risk will be the lowest. We introduce risk_i as a new value of r_{self} resulting from exporting to (hence trusting) e_i. Further, we define risk_0 to describe the lack of change to r_{self} (i.e. the case of not exporting anything and not trusting anyone). Thus, treating it as a conditional probability:

$$\mathrm{risk}_i = t_i() \times r_{\mathrm{self}}(c_{\mathrm{self}} - c_{\mathrm{export}} + c_{\mathrm{import}}^i)$$
$$+ (1 - t_i()) \times r_{self}(c_{\mathrm{self}} + c_{\mathrm{import}}^i)$$

$$\mathrm{risk}_0 = r_{\mathrm{self}}(c_{\mathrm{self}}).$$

Now the decision that the trustor has to make can be described as finding j such that $\mathrm{risk}_j = \min(\mathrm{risk}_i, i = 0, \ldots, n)$. If the outcome is that $j = 0$ then the trustor should not export anything at this time. Otherwise the trustor should export to e_i and trust this trustee.

9 Application

Before the model is verified through some experimentation, it should be verified by mental experiments: whether it explains defined cases and in this way whether it satisfies thesis presented earlier. This section discusses the application of the model twice: once through the lens of motivating cases, and then using a real-life example. In both cases, this section is looking for the answer whether the model is relevant to the case, whether it explains what happened and whether its explanation is somehow better than the one provided by existing models.

9.1 Application to motivating cases

There are features of this model that directly explain motivating cases introduced earlier in this paper. To better describe them, the diagram on Fig. 2. shows an exemplary function that links r_{self} and c_{self}, i.e. the level of risk that the trustor experiences because of its internal complexity. The function is likely to be monotonously increasing but not linear—the trustor may be not concerned much for as long as its complexity is low, and then the risk may rapidly increase as c_{self} is approaching c_{max}.

In reference to 'motivating cases' discussed earlier in this paper, we can observe the following.

1. If the trustor chooses one of the entities, all other things being equal, the more trustworthy one is chosen. This situation is the typical one discussed throughout the literature.

Let us consider the situation where the trustor is at c_{self}=0.7 $\times c_{max}$ and wants to export 0.25 of c_{max}. Let us assume that both entities will provide negligible import (c_{import}^1=0.1 and c_{import}^2=0.1) and are reasonably trustworthy with $t_1()$=0.8 and $t_2()$=0.7. For simplification, all functions that determine trustworthiness are constants. We have as follows (all values rounded to two decimal points):

$$risk_0 = 0.34$$
$$risk_1 = 0.23$$
$$risk_2 = 0.27.$$

In this situation, as expected, the trustor will choose the first entity, driven by the minimum risk it offers. It is also the more trustworthy entity, with its $t_1()$=0.8.
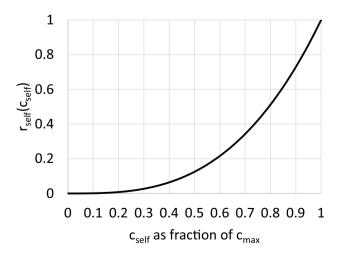

Fig. 2 An example of a risk to self-preservation

2. It is possible that the trustor chooses not to trust at all, specifically when both c_{self} and c_{export} are low. This describes situations where the trustor is comfortable with its current and possible future positions. If this is the case, it may decide not to engage despite having several potential entities to choose from.

Let us consider the situation where the trustor is at c_{self}=0.3 $\times c_{max}$ and wants to export only 0.1 of c_{max}. Let us assume that both entities will provide negligible import (c_{import}^1=0.1 and c_{import}^2=0.1) and are slightly less trustworthy with $t_1()$=0.6 and $t_2()$=0.5. Thus:

$$risk_0 = 0.03$$
$$risk_1 = 0.04$$
$$risk_2 = 0.05.$$

The model shows that in this case the trustor chooses not to trust anyone and handle the complexity all by itself. The main reason is that it is riskier to trust anyone than not to trust. Note, however, that this choice of not trusting is possible only when the trustor's complexity and risk are relatively low—the situation of a relative comfort.

3. The trustor can choose the less trustworthy entity over the more trustworthy one. This is one of situations that are not well explained by current models. It is where, despite having a choice of a more trustworthy entity, the trustor decides to trust the less trustworthy one.

This formalisation demonstrates that this is a rational choice (albeit possibly not in the best long-term interest of the trustor), provided that the less trustworthy entity adds less complexity to the trustor.

Let us consider the trustor at c_{self}=0.9 $\times c_{max}$ and wants to export 0.25 of c_{max}. Let us assume that both entities will require different imports (c_{import}^1=0.1 and c_{import}^2=0.2) and that the one with higher import is also more trustworthy with $t_1()$=0.7 and $t_2()$=0.8. Thus:

$$risk_0 = 0.73$$
$$risk_1 = 0.60$$
$$risk_2 = 0.75.$$

In this situation, the less trustworthy entity is chosen, as the overall risk is lower. This is due to the fact that for more trustworthy entity, the trustor has to bear the complexity of import that significantly increases its own risk. Note that this situation is more prevalent when the trustor has little capacity to spare, i.e. when the trustor is pressed to trust someone.

4. When trustor's own complexity c_{self} is low, the trustor is choosier, if it is high, the trustor is less choosy.

This has been already demonstrated by some of the cases above, e.g. the previous one. In general, certain comfort when it comes to decision-making and the ability not to export makes the trustor more selective, with preferences for engaging only with the best entities, and withholding if those entities are not available. Contrasting, being pressed for decision means that the trustor is more likely to make a decision that is maybe beneficial short-term but may be unfavourable long term.

5. The existence of a choice increases trust

Choice does not only mean that the trustor can make a single export, as it has been discussed above. It also means that the trustor can devise (or at least anticipate) the more complex strategy, considering some 'what-if' scenarios where the failure of one entity may lead to secondary choice of another entity.

From the perspective of a formalisation, the use of such a strategy means that the trustor may consider all elements as a kind of a collective entity and evaluate their collective trustworthiness (i.e. likelihood that at least one of them will accept the export to the satisfaction of the trustor).

Let us consider the scenario where the trustor is willing to export to the most trustworthy entity, and if it fails it is willing to accept the penalty and try the one that is less trustworthy. The calculation may now consider several cases, but for simplification we concentrate on a simple case of trusting e_1 and, if it fails, trusting e_2. The risk associated with this case will be described as risk_{1-2}.

Let us consider the trustor at $c_{\text{self}} = 0.7 \times c_{\text{max}}$ and wants to export 0.3 of c_{max}. Let us assume that both entities will require some imports ($c^1_{\text{import}} = 0.1$ and $c^2_{\text{import}} = 0.1$) and that one is more trustworthy with $t_1() = 0.8$ and $t_2() = 0.7$. Thus:

$$\text{risk}_0 = 0.34$$
$$\text{risk}_1 = 0.20$$
$$\text{risk}_2 = 0.24.$$

However, the risk of serial trusting is lower, with $\text{risk}_{1-2} = 0.17$. So, while the trustor may be content with the strategy of trusting e_1, it may be even more willing to try and trust e_1 and then e_2 knowing that such serial trusting bears significantly less risk to its self-preservation.

9.2 Real-life considerations

The real-life considerations, described in this section, focus on the question of quantification and reasoning. Indeed, the formalism of the model requires that both the complexity and the risk function should be somehow quantified.

It has been already mentioned that the complexity discussed here cannot be quantified simply by counting the number of communications that the system consists of. The more appropriate metric would come by an analogy to the content of a large database, where the ability to process data gradually diminishes as the database fills up with data. If there is a single metric, then probably the delay is response which is the easily observable one, while the subjective perception of the inability to cope with information overload is the one that can be captured by surveys (Kerr 1973).

The overall shape of the risk function is likely to depend on the nature of the trustor. It will be driven by the subjective perception of risk in case of individuals, but it is more likely to be based on risk-management processes for organisations.

Individual's risk function is based (among others) on works of Douglas (2013). It highlights the subjective perception of immunity to risk followed by the sharp increase in the perceived risk when situations are particularly severe. Perception of risk can be quantified using various subjective methods (Slovic 1992).

The independent variable of the risk function is the complexity. Initial thoughts about the link between risk, complexity and self-preservation came from Luhmann (2005), but several risk-management processes have a similar view that the complexity of the situation increases risk (Chapple et al. 2018). It is unlikely that the complexity is the only contributor to risk, but it is likely that it increases the perception of risk.

Finally, there is a question whether all real-life situations are resolved by reason, and it is definitely not true (Kahneman 2013). It is not even always possible or desired to employ the whole analysis presented here, where simple causation may be satisfactory. This, however, should not be taken against the model, as it does not attempt to explain the whole human behaviour, only to provide a computational approximation of it.

10 Conclusions

Trust, as usually discussed throughout the literature, is associated with a choice, thus ignoring cases where choice seems to be limited or non-existent and where trust is not granted despite having a choice. This paper formalises decision to trust on the basis of the theory of social systems to demonstrate that such formalisation extends into situations where trust defies choice. It shows that self-preservation can be the driving force of such behaviour.

The formalisation explains several situations that may seem irrational: trusting untrustworthy entities, trusting monopolies, resorting to trust in abstract systems rather than in actual entities or resolving not to trust at all. Thus, it allows to use the single mechanism of computational trust to describe the wide range of experiences.

The paper demonstrates that trusting in situations of no or limited choice is rationally explainable, as it preserves the integrity of self. That is, under some circumstances, it is better (at least short-term) for the trustor to trust someone who is not suitable than not to trust at all, as not trusting may threaten the very existence of the trustor. Conversely, it may be better for the trustor not to trust anyone than engage in a relationship that will threaten its self-preservation.

The formalisation interprets the act of trusting as exporting complexity from the trustor to a trustee, in exchange for importing some necessary overhead. Thus, trusting is not 'free of charge' for the trustor, as it is definitely not for the trustee. The introduction of the unified metric of cost, in a form of complexity, allows to explain some of the phenomena of trust with no or limited choice.

Further, the formalisation demonstrates that there is a need for only one, relatively simple, model that covers both situations of comfortably choosing the one to trust and situations where the trustor decides under duress, with no or little choice or where it withdraws from trusting. That is, there are no special mechanisms at work that emerges when choices diminish.

The main objective of this model is to enhance the domain of computational trust. Computational trust is used in technology in two distinctively different ways. First, it facilitates the development of automated recommendation systems (Golbeck 2009), a form of decision-support systems or 'trust advisors' (Cofta and Crane 2003). The quality of those systems depends not only on the availability of data, but first and foremost on using the reasoning scheme that closely resemble human thinking of trust and trustworthiness.

For such systems, the proposed model offers a benefit of incorporating aspects of decision-making that are currently ignored by existing solutions. It may be debatable, how useful this model may be for typical recommendations of books, music or other consumer goods, but for more complex cases, e.g. from the area of health, personal life etc., the inclusion of the self-preservation aspect may be beneficial.

The other way trust is incorporated into technological solutions is by treating it as a metaphor, and using to describe the operation of purely technical solutions. That is, instead of simulating all aspects of human trust, technology developers pick only relevant aspects while using the phrase 'trust' in a more liberal way. An example may be the trust-based routing where the node in the network makes decision on forwarding the packet on the basis of previous experience with other nodes, not on the basis of fixed routing tables.

The notion of self-preservation may not seem to be applicable to solutions such as routing. However, there are aspects, e.g. of service composition (Chang et al. 2006) where the node may be in a position of limited choice, and may be looking at the choice of the only provider of an inferior service, or stopping delivering its own service. In such situations this model may provide an additional insight. Similarly, agent-based systems (Wierzbicki 2010), may incorporate this model into algorithms of trusting agents.

The model may have implications beyond computational trust, as it lead to a better understanding of some behaviour associated with dealing with technology, specifically with monopolistic technology providers. This may lead to the development of new architectures and products, as well as new procedures that elicit and reinforce trust between customers and monopolistic providers.

This proposition requires further development, verification and integration. There are some outstanding questions such as the nature of complexity, the way the trustor can combat the complexity by itself, etc. Similarly, the nature of the risk function (both trustor's own and those that trustees represent to the trustor) requires empirical verification. They will be the subject of further works.

References

Beckert J (2005) Trust and the performative construction of markets. Max Planck Institute for the Study of Societies Discussion Paper 05/8

Bohnet I, Zeckerhauser Z (2004) Trust, risk and betrayal. J Econ Behav Organ 55:467–484

Borch K (1967) The theory of risk. J R Stat Soc Series B (Methodological) 29(3):432–467

Butler JV, Giuliano P, Guiso L (2012) Trust and cheating. NBER Working Paper Series. Working Paper 18509. http://www.nber.org/papers/w18509. Accessed 20 Dec 2018

Castelfranchi C, Falcone R (2000) Trust is much more than subjective probability: Mental components and sources of trust. In: Proc. of the 33rd Hawaii Int. Conf. on System Sciences (HICSS2000). Vol. 6, 2000. http://www.istc.cnr.it/T3/download/Trust-more-than-probability.pdf. Accessed 11 Mar 2005

Chang E, Dillion T, Hussain FK (2006) Trust and reputation for service-oriented environments: technologies for building business intelligence and consumer confidence. Wiley, Chichester, UK

Chapple M, Stewart JM, Gibson D (2018) CISSP official study guide. Wiley, Indianapolis. ISBN: 978-1119475934

Cofta P (2007) Trust, complexity and control: confidence in a convergent world. Wiley, Chichester, UK. https://doi.org/10.1002/9780470517857 (**9780470061305**)

Cofta P, Crane S (2003) Towards the intimate trust advisor. In: Proc. of the First Int. Conf., iTrust, 2003, Lect. Notes in Comp. Sci. 2692

Coleman JS (1982) Systems of TRust. Angewandte Sozialforschung 10:277–300

Deutsch M (1958) Trust and suspicion. J Confl Resolut 2(4):265–279. https://doi.org/10.1177/002200275800200401

Douglas M (2013) Risk and acceptability. Routledge, London. https://doi.org/10.4324/9780203708781 (**9780203708781**)

Festinger L (1957) A theory of cognitive dissonance. Stanford University Press, Stanford

Galli G, Nardin G (2003) Choice under uncertainty: the role of trust, 19th annual IMP conference, 4th-6th September. Lugano, Switzerland

Gambetta D (2000) Can We Trust Trust? In: Gambetta S (ed) Trust: making and breaking cooperative relations, electronic edition. Department of Sociology, University of Oxford, chapter 13, pp. 213-237. http://www.sociology.ox.ac.uk/papers/gambetta213-237. Accessed 27 Aug 2004

Giddens A (1990) The consequences of modernity. Polity Press, Cambridge. ISBN 978-0-7456-0923-2

Giddens A (1991) Modernity and self-identity. Self and society in the late modern age. Polity Press, Cambridge, UK (0-7456-0932-5)

Golbeck J (2009) Computing with social trust. Springer, London. https://doi.org/10.1007/978-1-84800-356-9. ISBN 978-1-84800-355-2

Hardin R (1991) Trusting persons, trusting institutions. In: Zeckhauser RJ et al (eds) Strategy and choice. MIT Press, Cambridge, Mass

Hardin R (2002) Trust and trustworthiness. Russel Sage Foundation, New York. ISBN: 0871543427

Harvey JS (2013) You can have your trust and calculativeness, too: uncertainty, trustworthiness and the Williamson Thesis. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2351870 or https://ssrn.com/abstract=2351870. Accessed 15 Dec 2018

Jøsang A, Presti SL (2004) Analysing the relationship between risk and trust. In: Jensen C, Poslad S, Dimitrakos T (eds) Trust management. iTrust 2004. Lecture notes in computer science, vol 2995. Springer, Berlin. https://doi.org/10.1007/978-3-540-24747-0_11

Jøsang A, Ismail R, Boyd C (2007) A survey of trust and reputation systems for online service provision. Decis Supp Syst 43(2):618–644. https://doi.org/10.1016/j.dss.2005.05.019

Kahneman D (2013) Thinking, fast and slow. Farrar, Straus and Giroux, New York. ISBN: 978-03745335577

Kao AC, Green DC, Davis NA, Koplan JP, Cleary PD (2001) Patients' trust in their physicians: effects of choice, continuity, and payment method. J Gen Intern Med. https://doi.org/10.1046/j.1525-1497.1998.00204.x

Karni E, Schmeidler D (1986) Self-preservation as a foundation of rational behavior under risk. J Econ Behav Organ 7(1):71–81

Kasten L (2018) Trustful behaviour is meaningful behaviour: implications for theory on identification-based trusting relations. J Trust Res. https://doi.org/10.1080/21515581.2018.1479967

Kerr B (1973) Processing demands during mental operations. Mem Cogn 1:401. https://doi.org/10.3758/BF03208899

Klausen E, Passman RH (2007) Pretend companions (imaginary playmates): the emergence of a field. J Genet Psychol 167(4):349–364. https://doi.org/10.3200/gntp.167.4.349-364 (PMID 17645227)

Lewicki RJ, Bunker BB (1995) Trust in relationships: a model of development and decline. In: Bunker BB, Rubin (eds) Conflict, cooperation and justice: essays inspired by works of Morton Deutsch. Jossey-Bass, San Francisco, pp 133–173

Luhmann N (1979) Trust and power. Wiley, Chichester, UK

Luhmann N (1988) Familiarity, confidence, trust: problems and alternatives. In: Gambetta D (ed) Trust: making and breaking cooperative relations. Blackwell, Oxford

Luhmann N (1995) Social systems. Stanford, California

Luhmann N (2005) Risk: a sociological theory. Aldine Transaction. ISBN: 978-0202307640

Marsh S (1994) "Formalizing trust as a computational concept". PhD thesis, University of Stirling, Department of Computer Science and Mathematics

Mather M, Shafir E, Johnson MK (2000) Misremembrance of options past: source monitoring and choice. Psychol Sci J 11(2):132–138. https://doi.org/10.1111/1467-9280.00228

Mayer RC, Davis JH, Schoorman FD (1995) An integrative model of organizational trust. Acad Manag Rev 20(3):709–734. https://doi.org/10.5465/amr.1995.9508080335

McKneally MF, Ignagani E, Martin DK, D'Cruz J (2004) The leap to trust: perspective of cholecystostomy patients on informed decision making and consent. J Am Coll Surg 199(1):51–57

McKnight DH, Chervany NL (2001) Conceptualizing trust: a typology and E-commerce customer relationships model. In: Proc. of the 34th Hawaii Int. Conf. on system sciences

McKnight DH, Kacmar C, Choudhury V (2003) Whoops… did I use the wrong concept to predict e-commerce trust? Modeling the risk-related effects of trust versus distrust concepts. In: Proceedings of the 36th Annual Hawaii international conference on system sciences, 6–9 Jan. 2003. https://doi.org/10.1109/hicss.2003.1174393

Möllering G (2006) Trust: reason, routine, reflexivity. Elsevier, Amsterdam. ISBN 978-0-08-044855-8

Möllering G (2014) Trust, calculativeness, and relationships: a special issue 20 years after Williamson's warning. J Trust Res 4(1):1–21

Nickel PJ, Vaesen K (2012) Trust and Risk. In: Roeser S et al (eds) Handbook of risk theory. Epistemology, decision theory, ethics, and social implications of risk. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-1433-5. ISBN 978-94-007-1432-8

Resnick P (2006) The value of reputation on eBay: a controlled experiment. Exp Economics 9(2): 79–101. http://www.si.umich.edu/~presnick/papers/postcards/PostcardsFinalPrePub.pdf . Accessed 25 Apr 2009

Sitkin SB, Pablo AL (1992) Reconceptualizing the determinants of risk behavior. Acad Manag Rev 17:9–38

Slovic P (1992) Perceptions of risk: reflections on the psychometric paradigm. In: Golding D, Krimsky S (eds) Theories of risk. Praeger, Westport

Solomon RC, Flores F (2003) Building trust in business, politics, relationships and life. Oxford University Press, New York

Tan Y-H, Thoen W (2000) Formal aspects of a generic model of trust for electronic commerce. In: Proc. of the 33rd Hawaii Int. Conf. on system sciences

Vanderhaegen F (2017) Towards increased systems resilience: new challenges based on dissonance control for human reliability in cyber-physical & human systems. Annu Rev Control 44:316–322

von Neumann J, Morgenstern O (1953) Theory of games and economic behavior. Princeton University Press, Princeton

Ward A, Smith J (2003) Trust and mistrust. Radical risk strategies in business relationships. Wiley, Chichester. ISBN: 978-0470853184

Wierzbicki A (2010) Trust and fairness on open, distributed systems. In: Studies in computational intelligence, vol. 298. Springer. https://doi.org/10.1007/978-3-642-13451-7 (ISBN 978-3-642-13450-0)

Williamson O E (1993) Calculativeness, Trust, and Economic Organization. The Journal of Law & Economics, Vol. 36, No. 1, Part 2, John M. Olin Centennial Conference in Law and Economics at the University of Chicago (Apr., 1993), pp. 453-486