# PlayGround 2.0: Simulating behavior decisions with trust and control computations

Etienne Cartolano, Antonio Mauro Saraiva

Depto. de Engenharia de Computação e Sistemas Digitais
Escola Politécnica da Universidade de São Paulo
São Paulo, Brasil
cartolano@usp.br, saraiva@usp.br

Robert D. Stevenson

Dept. of Biology
University of Massachusetts - Boston
Boston, USA
robert.stevenson@umb.edu

Abstract—Trust fosters cooperation in social interactions. It is widely studied in applied and academic disciplines. However, trust is a hard concept to define. Network analysis of the Web of Science citation databases shows it has multiple meanings and applications even within Computer Science. To date general models of trust are essentially conceptual, whereas workable proposals focus on special situations. In an effort to make trust simulations more generalist and applicable, we implemented the PlayGround model (Marsh, 1994) and modified it to incorporate a broader concept of trust and notions of control. Results with new model, that consider competence as a trustworthiness component, and use control to mitigate the risks, were similar to original ones from Marsh's, although the conceptual differences are significant to encompasses the many meanings of trust.

Keywords—trust model, simulation, control, trust meanings

#### I. INTRODUCTION

Trust is a social phenomenon experienced daily by people [1] and it is an inherent building block of our society [2]. We trust that drivers will respect the law and stop their cars before the crosswalk, that our national team will be champions, and that government agencies will be able to control new pandemics. However, there is a risk that they do not behave as we expect. Drivers may be texting and not see the signal to stop, our best player may go to a party before the championship game and get tired, or the government agencies may wait to act against new diseases because of political pressures. Generally speaking, we trust when engaging in a relationship. There is always a risk associated in a trust situation that others may not behave according to our expectations [3]. If there is no risk, then there is no reason to trust or distrust, because others will meet our expected.

The decision to engage as a trustor in a trust situation depends on factors such as the historic trustworthiness of trustee (the counterpart), their competence, the risks mentioned above, and the utility and importance of situation [4]. The connections among these factors are natural to us. For instance, it's very important to avoid being hit by a car, so that we will cross a busy road only if we expect an on-coming driver to stop their car. When our perception about a trustee's competence and interest to behave as we expect are high, and our perception about the risks are low, our confidence to engage in a trust situation will be high.

Depending on forms of trust for making decisions is not the whole story. It is possible to use instruments such as contracts, laws, inspections, and social rules, to mitigate risks, and increase confidence. As the Russian proverb made famous by Reagan says, "Trust, but verify" [5]. Trust and control play complimentary roles in the basis for confidence in relationships [6]. For instance, we can verify if the driver is paying attention by signalizing each other, or we can take the crosswalk near a crossing guard who is expected to regulate driver behavior. These actions are not a guarantee of expected behavior, because they intend to impact a driver's actions in the future. However, they act to mitigate the associated risks that driver will not stop and give us more confidence to engage in the trust situation of crossing the street in front of the car.

Despite the advantages of verification and control, their application has limits. For instance, when the bus driver takes our children to school, we can verify the company's and driver's experience, apply tests of quality, and perhaps even watch with security cameras, but we cannot be at every highway intersection. Generally speaking, humans are unable to face the complexities of the world without resorting to trust [7]. This means the Russian proverb: "Trust, but verify" does not tell the whole story. A revised version of the proverb might read "Trust, but verify... if resources allow".

In this work, we build a generalized model of behavioral choice that incorporates both trust and control. We begin by showing that trust is used to describe many differ kinds of situations within Computer Science (Section II). A review of the trust literature identifies Marsh's [4] work as a good starting point for computational trust models (Section III). We describe Marsh's implementation of his formalism called PlayGround and replicate the original experiments (Section IV). Next we introduce Cofta's notion of the joint role of trust and control (Section V), and we modify the PlayGround simulator to accommodate the idea of control using Cofta's [6] formulation (Section VI). In the final section (VII) we compare the results from original PlayGround model with results from the new model.

#### II. THE MANY MEANINGS OF TRUST

As in our daily lives, trust is studied widely in applied and academic disciplines such as: Business, Philosophy, Law, Psychology, Computer Science, Sociology, and Biology. There

are dozens of definitions and models for trust [8] [9] [10] [11] [12], but no one author, definition, or model is authoritative. Each area tends to view trust from its own perspective [13], focusing on subconcepts to create its particularized version of trust [9].

To explore the many meanings and applications of trust across disciplines we conducted a co-occurence analysis [14] with articles written in English that contained the words "trust", "trustworthy", and "trustworthiness" in the abstract, title, or keywords. The Web of Science database [15] returned 88,292 articles from 48 areas supporting the statement that trust is used in many disciplines. Within Computer Science the search resulted in 10,675 articles, which were the basis for constructing a co-occurrence graph using the VOSviewer [16] (Fig. 1).

Computer Science articles grouped in three clusters. Cluster A encompasses models and implementations of trust for security and privacy applications (blue, left hand cluster). Actors may be people, but are also computers, algorithms, application, hardware, or transactions. Cluster B articles are close to trust between people in general systems, social networks, reputation, and recommendation systems (green, center left). Cluster C relates trust as a support to business on Internet, that involve companies, brands, transactions, and third-parts (red, right cluster). Other areas cited by Computer Science authors include cognitive sciences [17], management [11], sociology [7], and political sciences [3]. This exploratory analysis demonstrates that within computer science trust has many meanings and applications. The challenge for scientists has been to find common ground among the many uses of trust.

# III. GENERALIZING TRUST CONCEPTS

Because trust is central to cooperation in many domains [9] [3] [18], simulating its dynamics is widely important. There are

many applications for simulations in computer science (Fig. 1). In other disciplines, simulations can: predict team behavior for different challenges, or under different styles of leadership; reproduce the dynamics of group formation in societies of employees, students, animals, or artificial agents; or test trust for evolution of consumer goods preferences or species social structure. However, to build models of trust dynamics that are more widely applicable it is necessary to identify generalities about trust and how it works.

In a trust situation, the trustor expects the trustee to undertake a specific behavior. There is an inherent risk that the trustee may not behave as expected. The decision to trust (or distrust) the trustee, and being able to engage (or not) in a specific relationship, is result of the trustors' general capacity to take risks (= trustfulness), the assessment of trustee's interests and capacity to behavior as the trustor expects (= trustworthiness), and the instruments available to the trustor to enforce this behavior.

Among several efforts to model trust relationships [10], Marsh's thesis was one of the first computational models to simulate trust behavior. It is the starting point for our efforts to build a more generalized trust and control model. We begin by describing and replicating Marsh's efforts.

# IV. MARSH'S FORMALISM FOR TRUST

Marsh [4] developed a formalization and a computation model for trust based on the agent concept of Artificial Intelligence (AI). The key idea was to extent an AI agent's reasoning capabilities to reproduce trust behavior - that is to decide when and with whom to cooperate. In his model this decision depends on three kinds of trust (Basic, General, and Situational Trust), and contextual elements including Risk, Importance, Utility, and Competence.

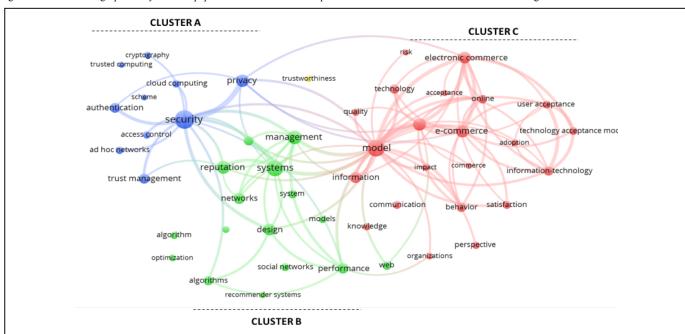


Fig. 1. Co-occurrence graph of keywords in papers related to trust in Computer Science Web of Science database made using VOSViewer.

# A. Marsh's Formalism and Model of Trust

To cooperate in a situation  $(\alpha)$ , Alice's (x) trust in Bob (y) must be greater them a threshold (eq. 1). This trust  $(T_x(y, \alpha), eq. 2)$  depends on the Importance  $(I_x(\alpha))$  and the Utility  $(U_{\mathbf{v}}(\alpha))$  of situation for Alice, that are respectively a rational economic and a subjective judgments of situation, and on the Estimated General Trust (EGT or  $\widehat{T_x(y)}$ ) of Alice in Bob. The General Trust (GT or  $T_{\mathbf{x}}(\mathbf{y})$ ) denotes how much an agent trusts another agent in general when the decision is being made without reference to a specific situation. For instance, Alice may say "I trust in Bob" based in an analyze of historic interactions with him. The definition of the EGT depends on Alice's memory span and her Disposition to Trust. Alice can use the highest past GT value in Bob in a similar situation (Optimistic disposition), the lowest value (Pessimistic), or a mean of values (Realistic, eq. 4) to define the EGT. A in eq. 4 represents the set of all situations similar to  $(\alpha)$  which Alice has experienced with Bob ([4], pg. 67). If Bob is unknown in a similar situation, GT past values in any situation are used to define the EGT. When there is not a past value for GT, then Alice's Basic Trust (BT) is used. BT derives from her past life experiences, and doesn't have a connection with other agents, a specific situation, or the environment. Simply, "Alice trusts". The Cooperation Threshold (eq. 3) is an estimative of how much Perceived Competence and EGT mitigate the Perceived Risk, multiplied by the Importance of the situation to Alice. The Perceived Risk ([4], pg. 71) and the Perceived Competence ([4], pg. 73) of Alice in Bob depend on the knowledge between them from previous interactions.

$$T_x(y,\alpha) > CT_x(\alpha) \implies Will\_Cooperate(x,y,\alpha)$$
 (1)

$$T_x(y,\alpha) = U_x(\alpha) \times I_x(\alpha) \times \widehat{T_x(y)}$$
 (2)

$$CT_{x}(y,\alpha) = \frac{Perceived\_Risk_{x}(\alpha)}{Perceived\_Competence_{x}(y,\alpha) + T_{x}(y)} \times I_{x}(\alpha)$$
(3)

$$\widehat{T_x(y)} = \frac{1}{|A|} \sum_{\alpha \in A} T_x(y) \tag{4}$$

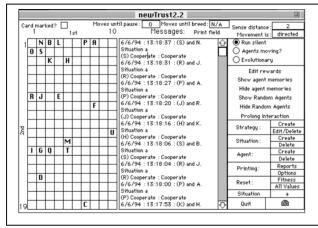
# B. The PlayGround

Marsh developed two implementations of his theory in chapter 7 [4]. The first considered trust as a strategy in an Axelrod's Tournament [19] and proved the feasibility of the formalism. The second, called the PlayGround, implemented the formalism more completely, by introducing some additional features to the tournament concept, aiming to create an environment closer to the real world.

The PlayGround was a graphical application (Fig. 2) developed in HyperCard that simulated interactions among agents. Agents moved around a cellular grid. An interaction occurred when an agent tried to move to a place that was already occupied. If occupied, agents played a round of the Iterated Prisoner's Dilemmas (IPD) [19] game in which they 'Cooperate' or 'Defect'. The payoff matrix varied according to the situation assigned to the interaction, which was randomly assigned, or defined as the same for all interactions. A generic payoff (T=>5, R=>3, P=>1, and

S=>0) was used for most experiments. At the end of all movements, the agent with highest score won.

Fig. 2. Screenshot of original PlayGround.



The direction of movement for all agents was chosen randomly along grid axes, except for trusting agents, which were able to choose a movement toward a trustworthy agent (with limitations described below). This choice was limited by a configurable 'range of vision' of the grid, equal for all directions, for example, two positions in each direction. Although the trusting agent chose a direction, the actual direction was determined by a weighted random choice: 50% for the desired direction, 10% in the opposite direction, and 20% for each of the two other directions. This directed movement was designed to reflect human behavior, in which an individual usually has access to a circle of trustworthy people allowing agents to form groups, and also to reflect external factors that make impossible the desired movement.

Other feature in which the PlayGround approximated real life was the adjustment of trust values after an interaction. These adjustments apply to Basic Trust, General Trust, and Perceived Competence. For instance, in the case of Alice trusting Bob, if both agents cooperated, then Alice's Basic Trust and General Trust increased 1% and 10%, respectively, and Alice's Perceived Competence of Bob was increase by 1% (Table I). Further details of implementations can be found in original work [4].

TABLE I. TRUST VALUES ADJUSTMENTS FOR ALICE

Alice & Bob actions	Basic Trust	General Trust	Perceived Competence	
Both Cooperate	+1%	+10%	+1%	
Cooperates and Defects	-1%	-10%	-1%	
Defects and Cooperates	5%	1%	+1%	
Both Defect	-5%	-10%	-1%	

# C. Replicating PlayGround's experiments

In this section we summarize the experiments conducted by Marsh using PlayGround. They were used as evidences of our correct replication of the original simulator in Section VII, and used also as a testbed to our proposal in Section VIII. Further discussion about the original experiments can be found in [4]. The first group of experiments was designed to test if trust could be learned by a trusting agent. Two trusting agents were forced to play an IPD. They had an unbound memory and they could not move around the grid. Agent A (realistic), that became C (optimistic) during experiments, had extremely low initial values of trust in all experiments and set up to distrust (defect) the counterpart, agent B, that became D, but was always a realistic agent, with same values, and remained cooperating all time. The speed of learning was measured as the number of interactions that agent A/C needed to trust and cooperates with agent B/D.

The variation of results between the experiments were directly connected with the way Estimated General Trust is calculated, and because the counterpart always cooperated. The speed of learning of a realistic agent depends on the number of past interactions used to calculate the EGT (eq. 4). The lower the number of past interactions remembered, the greater become EGT (because the counterpart always cooperates); Consequently, the Cooperation Threshold (eq. 3) became lower, and the speed of learning faster.

Using PlayGround 2.0 described below (Section VII), we replicate Marsh's original calculations of his Experiment  $\mathcal{A}$  on the speed of learning and compare Table II with his results on [4], pg. 161. The optimistic agent C starts to cooperate at the 13<sup>th</sup> interaction. Our output values are exactly the same as Marsh's, but the adjustment scores for the Basic and General Trust columns in original work are reversed until agent C starts to cooperate, then adjustments duplicate Marsh's table. We obtained exactly the same trust values for counterpart agent (D).

TABLE II. RESULTS OF EXPERIMENT  $\mathcal A$  REPLICATION

Int	Basic	General	Compe- tence	Situation	Threshold	Act
13	0.143668	0.090146	0.007212	0.620646	0.572641	C
12	0.136827	0.089253	0.007141	0.591093	0.600133	D
11	0.130311	0.088369	0.007070	0.562944	0.628908	D
10	0.124106	0.087494	0.007000	0.536138	0.659009	D
9	0.118196	0.086628	0.006931	0.510607	0.690498	D
8	0.112568	0.085770	0.006862	0.486294	0.723436	D
7	0.107208	0.084921	0.006794	0.463139	0.757881	D
6	0.102103	0.084080	0.006727	0.441085	0.793899	D
5	0.097241	0.083248	0.006660	0.420081	0.831561	D
4	0.092610	0.082424	0.006594	0.400075	0.870933	D
3	0.088200	0.081608	0.006529	0.381024	0.912075	D
2	0.084000	0.080800	0.006464	0.362880	0.955076	D
1	0.080000	0.080000	0.006400	0.345600	1.000000	D

In a further effort to validate Marsh's simulations, we reran his experiment  $\mathcal{F}$  on the speed and efficiency in group formation. The grid was populated with 6 Random, red, and 5 Optimistic, green, Realistic, yellow, and Pessimistic, gray agents. Agents were configured to move directed toward trustworthy agents using a range of vision of 2 cells. Initial random values were generated for Competence, Importance, Risk, and Basic Trust, considering the range of values from Experiment  $\mathcal{A}$ . Because the initial conditions are generated randomly, and the direct movement has a random

component, it is not possible to duplicate Marsh's results exactly. However, after being assigned randomly on the grid at the beginning of the experiment (Fig. 3a), agents were also grouped after 250 interactions (Fig. 3b) like in original experiment.

Fig. 3. Agents after 250 inteactions in original experiment  $\mathcal{F}$ .

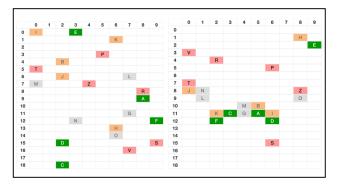


Table III presents the top and bottom five agents that held highest and lowest scores after 250 and 10,000 movements. Our results are different from Marsh's results, because none of our Random agents appears among first agents. Instead, Random agents are the among last in rank.

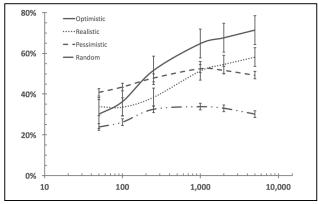
TABLE III. RANK OF AGENTS IN EXPERIMENT  $\mathcal{F}$  (COMPETENCE)

	After 250 Movements				After 10,000 Movements			
	Agent		Score	Int	Agent		Score	Int
1	P	Maria	456	143	R	Bob	13,340	5,579
2	R	Hardin	408	144	О	Charles	13,243	5,669
3	R	Kyle	354	123	О	Etienne	13,220	5,507
4	R	Ian	339	134	P	Neo	12,848	5,211
5	О	Fred	321	134	P	Oddin	12,567	5,129
17	О	Charles	214	127	RD	Silva	6,650	2,822
18	RD	Silva	201	65	RD	Rafa	6,242	2,777
19	RD	Peter	199	81	RD	Peter	6,218	2,788
20	RD	Tom	171	84	RD	Tom	5,938	2,767
21	RD	Zoom	139	60	RD	Zoom	5,669	2,667

This result is related to the number of interactions that agents have during their movements. An agent may need one or more movements to interact whit other agents. Trusting agents were more efficient in terms of interactions per movements, once they tend to direct their movement to more trustworthy agents.

After replicate the experiments, we proposed a new test for original model by analyzing successful cooperation in time. Initial random values were saved from Experience  $\mathcal{F}$  and used in five rounds. Optimistic, Realistic, Pessimistic, and Random values in Fig. 4 are means of the values of their agents in the five rounds. When simulations are done until 10,000 movements, clear differences among the strategies emerge. Optimistic and realistic agents have a mean of 71% and 60% of successful cooperation among all interactions and they increase successful cooperation consistently, while the pessimistic and random agents have lower means and show decreasing % of successful cooperation.

Fig. 4. Successful cooperation evolution in original Experiment  $\mathcal{F}$ .



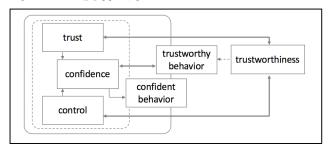
Overall, our efforts to reproduce Marsh's results were successful. A few model and implementation details were not explicit in Marsh's work. These are clarified in section VII. The last task is to modify Marsh's model to include new ideas about behavior choices.

#### V. COFTA'S FORMALISM FOR TRUST AND CONTROL

Cofta [6] aims to understand the confidence an agent has in behavioral choices. Confidence is presented as a belief that another will behave as expected, allowing people to engage in relationships by externalizing a set of actions (confident behavior). Cofta views confidence as a function of trust and control, defined as a "subjective probability of expectation that a certain desired event will happen, if the course of action is believed to depend on another agent". Confidence is assessed in a context perspective of time, agents, actions, and environment, that the author calls a transaction.

The general idea proposed by Cofta is that Alice (an intelligent agent) can be confident (and decide to behave in a cooperative way) either if she trusts Bob (another intelligent agent) or if she can control him (Fig. 5 from [6], pg. 14). Trust is equivalent to Alice's estimate about the internal intentions of Bob within the scope of a transaction (= trustworthiness). This estimate uses evidence about Bob's behavior divided in three categories: continuity, competence, and motivation ([6], pg. 39).

Fig. 5. Cofta's ([6], pg. 14) general model for behavioral choices.



One or more instruments of control (a policeman, bank's reputation, contracts) can be used to enforce a favorable action from Bob. The gap between Bob's trustworthiness

and his actions lies in the control element. Cofta identifies influence, knowledge, and reassurance as modes of control ([6], pg. 40). Each of these can improve Alice's judgement of Bob. Note there is a recursive nature of trust about the sources of evidence i.e. Can Alice trust this source of evidence about Bob's behavior? This recursivity is a clear example of our previous comments about the cost and capacity to verification all risks associated with trust in the Introduction.

# VI. PROPOSAL FOR A GENERAL MODEL OF TRUST AND CONTROL

Marsh and Cofta's have two complementary points of view about the cooperation behavior decisions. In a general perspective, the cooperation decision defined by Marsh (eq. 3) can be considered as the behavioral choice equivalent to Cofta's conceptual model (Fig. 5). Alice may cooperate if she has more trust than a cooperation threshold, or if she has evidences about the trustworthy behavior of the counterpart. However, although broader, Cofta's conceptual model does not offer a computational experiment that explicit tests how trust and control influent confidence; whereas Marsh's understanding of trust, although focused on competence, proposes a workable application for his formalism.

Our proposal encompasses these views in a general and workable model melding Cofta's more complete view of behavioral decisions, and Marsh's insights for a workable application. The main properties of this new model are as follows:

- Marsh's Cooperation Threshold (eq. 3) is still used as a basis for the decision to cooperate or not. If trust is greater than a threshold the agent will cooperate.
- As proposed by Cofta, Competence is considered as a category of evidence that is encompassed by trustworthiness, instead of an element apart from trust (eq. 3). This approach guarantees that other forms of evidence, such as motivation, for example, will have similar influence on the decision.
- Risk mitigation proposed by Marsh is improved by the control element proposed by Cofta. This means that the perceived risk in eq. 3 will be compensated not only by trust but also by the control Alice has over their counterpart, Bob.

These new proprieties require a reformulation of eq. 3 as:

$$CT_{x}(y,\alpha) = \frac{Perceived\_Risk_{x}(\alpha) \times Importance_{x}(\alpha)}{T_{x}(y) + Control_{x}(y,\alpha)}$$
 (5)

The range of values of Risk, Perceived Importance, and EGT remain the same used in the original PlayGround, between 0 and 1. Thus, the values of control are between 0 and 1 also, excluding the limits.

# VII. THE PLAYGROUND 2.0

PlayGround 2.0 is implemented in PHP. The simulator and the code are available online in [20]. Implementation details about types of agents, their movements, and

adjustment of trust values were collect in [4]. Listed below some important factors that were assumed from original simulator, or introduced by us:

- Movement: The original grid was 'flat', and an agent could not move to east if he was on the most eastern position, remaining in the same position. Now, it is possible configure a wrap field.
- Directed Movement: The trusting agent will consider only the first (and closest) agent that he sees in a direction to choose a movement. A simulated interaction using the generic situation is done with existent neighbors in four directions. The greatest difference between the Situational Trust and the Cooperation Threshold (ST-CT), even negative values, will define the chosen movement. In case of absence of neighbors the movement will be random.
- Interaction: When two trusting agents interact, both play the roles of trustor and trustee for the exactly same situation, because both need trust to decide the action, so that in an unique interaction there are two plays ("A trust B"and "B trust A") to same situation. Trust values are adjusted for trustors and trustees.
- Disposition to Trust: For a Realistic agent, before calculating Situational Trust, a mean of past GT is saved as a new GT value in memory to use in future calculations. After interaction, the adjusted value from the mean will be saved as another new GT value. This is not true for Optimistic and Pessimistic agents, for whom only the GT originated from adjustments are saved as new values.
- Values: Although they had formulae, or suggestions of calculations, Importance, Utility, Risk, Costs, and Perceived Competence, had subjective values in original simulator ([4], pg. 83). So that, we decided to use random values between the lowest and highest values used in Marsh's experiments. We adopted the same strategy for Basic Trust. General Trust started with null values. Zero and one (1) values cannot be accepted for Competence, Control and Perceived Competence.

The first goal of the new simulator was to replicate some experiments and results of original simulator which are described above in Section IV. Next we run modified proposal (eq. 5) encompassing trust and control.

#### VIII. RESULTS OF TRUST SIMULATIONS WITH CONTROL

After implement the formulae (eq. 5), replacing original formulae to Cooperation Threshold (eq. 3) we reran the same experiments  $\mathcal{A}$  and  $\mathcal{F}$ . Agent C still started to cooperate at  $13^{th}$  interaction with relative low values. Fig. 6 show a same behavior of grouping between agents after 250 interactions.

Fig. 6. Agents after 250 interactions in experiment  $\mathcal{F}$  (control).

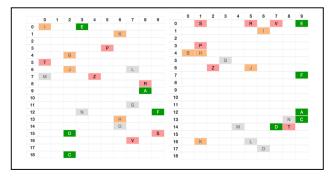


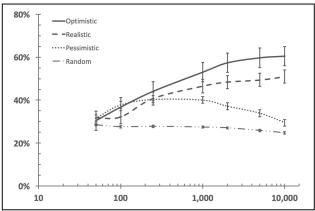
TABLE IV. RANK OF AGENTS IN EXPERIMENT  $\mathcal{F}$  (CONTROL)

	After 250 Movements				After 10,000 Movements			
	Agent		Score	Int	Agent		Score	Int
1	P	Maria	369	123	R	Bob	12,643	5,901
2	R	Hardin	367	135	P	Leo	12,387	5,476
3	R	Ian	341	140	О	Fred	11,920	5,678
4	R	Jean	332	135	P	George	11,749	5,299
5	P	Oddin	325	116	О	Etienne	11,349	5,566
17	RD	Silva	175	83	RD	Silva	5,846	3,047
18	RD	Tom	153	78	RD	Vick	5,743	2,988
19	RD	Rafa	151	65	RD	Rafa	5,544	2,782
20	RD	Peter	126	67	RD	Peter	5,443	2,868
21	RD	Zoom	120	51	RD	Tom	5,344	2,709

Table IV don't show changes related to Random agents, they remain out of top five scores and populate the last positions. Agents interacted at same level of model with competence, but score are lower.

The evolution of successful cooperation shows the same pattern in which Optimistic and Realistic agents has a consistent grow, but start to establish the results after 5,000 movements. Pessimistic agents have a low initial value and start to decrease soon, after 1,000 movements. Random agents have an average more consistent.

Fig. 7. Successful cooperation evolution in Experiment  $\mathcal{F}$  (control).



# IX. CONCLUSIONS

As a main contribution, this work introduced a generic trust model that encompasses many meanings of the word "trust" in a simulator of behavioral decisions. A brief analysis of common situations in our daily lives lists many uses and applications for the word "trust". The co-occurrence analysis showed that this situation is reflected by computer science with trust being analyzed in many different situations. In this scenario, Marsh's formalism and model was extended with Cofta's notions of trust and control to build a more generalized workable model for trust.

Our proposal was validated by using an implementation of Playground, a simulator created in Marsh's thesis. Our efforts to reproduce it were generally successful. Model and implementation details that were not explicit in original work are listed for future implementations.

Results with new model, that consider competence as a trustworthiness component, and use control to mitigate the risks, were similar to original ones from Marsh's. This fact is due to the mathematical similarity between eq. 3 and eq. 5, although the conceptual differences are significant. Competence and Control have the same range of values, and were initiated with random values in our experiences, so that at limit, they behave in a similar way. However, in real life simulations, control and competence will get different importance and values according to situations.

# X. FUTURE

Considering the many meanings of trust and the many disciplines in which it has been identified as a critical factor, more attempts to formulate models that can address a broad arrange of trust and decision making situations are needed. A second clear challenge are experimental studies to tease apart the components of trust models and to get quantitative estimates of parameter values. Other strategies can be used in experiments. Recent advances in the neurochemistry of trust combined with traditional behavioral studies offer exciting opportunities [21].

# XI. ACKNOWLEDGMENT

Authors thank the program Science without Borders of Brazilian Federal Government for the scholarship granted to the first author during his visit at University of Massachusetts - Boston.

#### XII. REFERENCES

- [1] A. Baier, "Trust and Antitrust," Ethics, vol. 96, no. 2, pp. 231–260, 1986.
- [2] F. Fukuyama, Trust: The Social Virtues and The Creation of Prosperity. New York: Free Press, 1996.
- [3] R. Hardin, Trust and Trustworthiness. New York: Russell Sage Foundation, 2002.
- [4] S. P. Marsh, "Formalising Trust as a Computational Concept," University of Stirling, 1994.
- [5] "Wikipedia: Trust, but verify," Wikipedia, 2016. [Online]. Available: https://en.wikipedia.org/wiki/Trust,\_but\_verify. [Accessed: 14-Aug-2016].
- [6] P. Cofta, Trust, Complexity and Control: Confidence in a Convergent World. Chichester: John Wiley & Sons Ltd, 2007.
- [7] N. Luhmann, Trust and Power. Chichester: John Wiley & Sons Ltd, 1979.
- [8] D. H. MckNight and N. L. Chervany, "The Meanings of Trust," Minneapolis, 1996.
- [9] P. C. Bauer, "Three Essays on the Concept of Trust and its Foundations," University of Bern, 2015.
- [10] J.-H. Cho, K. Chan, and S. Adali, "A Survey on Trust Modeling," ACM Comput. Surv., vol. 48, no. 2, pp. 1–40, 2015.
- [11] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An Integrative Model of Organizational Trust," Acad. Manag. Rev., vol. 20, no. 3, pp. 709–734, 1995.
- [12] A. Josang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," Decis. Support Syst., vol. 43, no. 2, pp. 618–644, 2007.
- [13] D. H. Mcknight and N. L. Chervany, "What Trust Means in E-Commerce Customer Relationships: An Interdisciplinary Conceptual Typology.," Int. J. Electron. Commer., vol. 6, no. 2, pp. 35–59, 2001.
- [14] L. Leydesdorff and L. Vaughan, "Co-occurrence matrices and their applications in information science: Extending ACA to the web environment," J. Am. Soc. Inf. Sci. Technol., vol. 57, no. 12, pp. 1616–1628, 2006.
- [15] Thomson Reuters, "Web of Science," 2016. [Online]. Available: https://webofknowledge.com/. [Accessed: 03-May-2016].
- [16] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," Scientometrics, vol. 84, no. 2, pp. 523–538, 2010.
- [17] C. Castelfranchi and R. Falcone, "Trust and control: A dialectic link," Appl. Artif. Intell., vol. 14, no. 8, pp. 799–823, 2000.
- [18] D. H. McKnight, V. Choudhury, and C. Kacmar, "Developing and validating trust measures for e-commerce: An integrative typology," Inf. Syst. Res., vol. 13, no. 3, pp. 334–359, 2002.
- [19] R. Axelrod, The Evolution of Cooperation. New York: Basic Books, 1984.
- [20] E. Cartolano, "PlayGround 2.0," 2016. [Online]. Available: www.trust.cartolano.com.br. [Accessed: 14-Aug-2016].
- [21] M. Kosfeld, M. Heinrichs, P. J. Zak, U. Fischbacher, and E. Fehr, "Oxytocin increases trust in humans.," Nature, vol. 435, no. 7042, pp. 673–677, 2005.