EPISTEME NS, VOL. 42, 2022, pp. 1-35

Luz Marina Barreto Guerra

RATIONAL CONSENSUS AND DELIBERATIVE DEMOCRACY IN COMPLEX SOCIETIES DOMINATED BY ONLINE SOCIAL MEDIA INTERACTIONS

Abstract: The purpose of my article is to explore the effects that complex communicative interactions in online social settings may have on our understanding of truth and validity, as these concepts have been defined in modern practical philosophy. For the latter, truth and the validity of democratic institutions can be assessed by a rational consensus that is, firstly, potentially universal for all possible rational beings or interlocutors, and, secondly, that can be interpreted as a reflective equilibrium achieved amidst an argumentative community of speakers. I aim at contrasting both intuitions with what *really* happen when rational agents confront their opinion in a real space of argumentative debate of complex nature, with possibly millions of interlocutors in principle accessible to the speaker: the public agora represented now by online social networks. I argue that, perhaps due to restrictions of *physical* nature, when social media corporations model the space of online conversations there could be a critical threshold where an opinion, as true as it may be, cannot be universalized, and where all "reflective equilibria", as understood by John Rawls, are unstable. In paragraph 7 of my paper I develop an argument as to why this must be so, which depends on the effect of complexity models for 1) the goal of reaching a suitable choice or balance between the variance of attributes within a model and the bias represented by truth (the problem posited by the so-called "bias-variance tradeoff"), and 2) for the goal of modeling true or valid semantic meanings in multidimensional frameworks. I conclude that both restrictions, the bias-variance tradeoff and the so-called "curse of dimensionality", maybe of true physical nature rather than a problem of algorithm design, do tell us something important about our philosophical understanding of truth and validity of statements and judgments about the world. They challenge the ancient philosophical intuition that truth can be, *ceteris paribus*, shared universally throughout a space of rational interlocutors,

Recibido 02-09-22 ☼ Aceptado 19-09-22

because, for the first time in history, it is possible to see that the clause "all things being equal" cannot be possibly fulfilled in a complex society. Finally, this would mean that universality of consensus is neither a necessary nor a sufficient condition for the establishment of truth of a statement.

Keywords: Rational consensus, deliberative democracy, online social networks, truth, deliberative democracy, bias-variance tradeoff and semantic multidimensionality.

CONSENSO RACIONAL Y DEMOCRACIA DELIBERATIVA EN SOCIEDADES COMPLEJAS DOMINADAS POR INTERACCIONES EN REDES SOCIALES

Resumen: El propósito de mi artículo es el de explorar los efectos que interacciones complejas de índole comunicativa pudieran tener sobre nuestra comprensión de la verdad y validez, tal y como ambas han sido definidas por la filosofía práctica moderna. Para esta última, la verdad o la validez de instituciones democráticas pueden ser establecidas por un consenso racional que es, en primer lugar, potencialmente universal para todos los posibles seres o interlocutores racionales y que, en segundo lugar, puede ser interpretado como un equilibrio reflexivo alcanzado por una comunidad argumentativa de hablantes. Mi propósito es contrastar ambas intuiciones con lo que pasa *realmente* cuando agentes racionales confrontan sus opiniones en un espacio de debate argumentativo complejo, en donde sería en principio posible para un hablante acceder a una audiencia de millones de interlocutores: el ágora pública representada hoy en día por las redes sociales online. Sugiero que, debido a restricciones de naturaleza física, cuando se intenta modelar el espacio de las conversaciones que ocurren online, pudiera haber un umbral crítico más allá del cual una opinión, por más verdadera que sea, no puede ser universalizada, y donde todos los equilibrios reflexivos, como los definió John Rawls en 1971, son inestables. En el parágrafo 7 de mi artículo desarrollo un argumento de por qué esto debe ser así, un argumento que depende 1) de las dificultades que entraña encontrar un balance adecuado entre la varianza de los atributos de un modelo y el sesgo que representa la verdad (un problema conocido como el "intercambio sesgo-varianza"), así como 2) de las dificultades para modelar significados semánticos verdaderos y válidos en espacios multidimensionales, cuando se tienen delante contextos semánticos complejos. Concluyo que ambas restricciones, el intercambio sesgo-varianza y la "maldición de la multidi-

mensionalidad", que representan probablemente restricciones de índole física más que un problema en el diseño de algoritmos, nos dicen algo verdaderamente importante acerca de nuestra comprensión filosófica de los conceptos de verdad y validez de enunciados y juicios de valor sobre el mundo. Ambos desafían la antigua convicción filosófica de acuerdo con la cual la verdad, *ceteris paribus*, puede ser compartida universalmente en un espacio de interlocutores racionales, dado que por primera vez en la historia es posible observar que la cláusula "si todas las condiciones siguen igual" nunca podrá ser satisfecha en una sociedad compleja. Esto significaría, finalmente, que la condición de la universalidad de los consensos no es ni necesaria ni suficiente para el establecimiento de la verdad de un enunciado.

Palabras clave: Consenso racional, democracia deliberativa, redes sociales online, verdad, democracia deliberativa, intercambio sesgo-varianza y multidimensionalidad semántica.

1. Our concept of reason

The question addressed in my paper is how do online social platforms affect the emergence of rational consensuses, usually considered of universal scope, between individuals communicating with each other and exchanging opinions in complex online settings and massive social networks. The question is relevant for us philosophers because throughout most of the history of philosophy a *universal* agreement on statements and value judgments has been considered a decisive mark of their truth or validity, as well as an expression of the robustness of human reasoning taking place inside argumentative processes. However, the behavior of massive online social networks seems to suggest that beyond some critical threshold, of some numerical value, a universal consensus might be not attainable among fragmented groups of users or individuals spread all over the massive space of online social media. This has decisive consequences for what is considered true or truly relevant within complex societies.

The legitimacy of public and political institutions has been traditionally considered to rest upon the constitution of a common political will that attempts to harmonize different individual goals and preferences under a common aim. For that, we do need shared conventions and intuitions, and these

presuppose communicating and "negotiating" different conceptions of the objective and the intersubjective world, hoping that these will eventually converge if all arguments are heard and freely uttered.

The thesis advanced in my paper is that in complex communicative settings, for reasons that need to be explained, our philosophical and cognitive intuitions collide with processes of physical nature that are, in a sense, insurmountable unless democratic correctives, enforced by the rule of law, are put into place.

Rational is usually predicated on a *true* statement or a *valid* judgment that is, in principle, *always* susceptible to be shared by others[1].

However, when we speak about a rational consensus, our tendency is to see it only referred to objective truth, understood as the result of what Kant defined as the theoretical use of reason. But, in Kantian vein, we may distinguish also between what it is "evidently" true for any individual having an experience (like "it is raining now"), and what needs to be *constructed* as deserving to be considered true by a community of argumentation (like to believe in the freedom of the will, or the existence of God).

Indeed, the constitution of rational consensuses cannot be understood without considering the question of how a statement or a normative judgment can have authority over me as a rational individual. This is a condition defined by Christine Korsgaard as "transparency", which states that we must know before taking action why a reason should motivate us, why do we think an action is justified or "make sense".[2]

But transparency also seems to be a demand for *all* kinds of rational statements, because reason presupposes an autonomous reflection on the arguments being considered by a rational being: to affirm the rationality of a

1 This is typically a Kantian conception of truth, developed in contemporary times by Karl Popper and Jürgen Habermas. See, in particular, their dispute in the fa mous colloquium published as *Der Positivismusstreit in der deutschen Soziologie* in 1969. Another important rendition of the Kantian concept of truth as universality can be found in O'Neill, O., *Constructions of Reason. Exploration of Kant's Practical Philosophy,*Cambridge University Press, 1989.

2 Cf. Korsgaard, C., *The Sources of Normativity*, Cambridge University Press, 1996, p.17.

statement or judgment implies to project it on to a universal audience *whose agreement with us we want in support, or in confirmation, of our self-reflection* that Korsgaard deemed to be also "self-conscious", as if in doubt until someone else agrees with us.

As a matter of fact, if the success of online social networks has clearly demonstrated something is that we crave "likes", that is, others to confirm what we believe to be true. We really need and want to agree with others, to be in consensus with them. Therefore, the Kantian stipulation, by which we accept as true or valid those statements and judgments that can be universally shared, seems to stand as a good definition of what we may consider rational or rationally grounded.

The mechanisms through which we can reach an agreement with another human being, feeling, at the same time, that we *really* agree with another when we interrogate ourselves and reflect in silence about what is being argued, define one of the most important and ancient problems of philosophical inquiry: why and how can we agree about something in the world, and can we say that this consensus is rational.

For instance, Aristotle's idea that basic intuitions or categories constituted a fundamental fabric of basic or general concepts, without which it would be impossible for two persons even to speak intelligibly to each other, is at the basis of his concept of science, *episteme*, understood as the ordered articulation of concepts getting from axioms, or undisputed intuitions about the world that everyone purportedly shared, to more empirical ones that needed to be explained or constructed.

This paradigm of sound knowledge reigned supreme for many centuries. Even at the turn of the 1900-century, many philosophers were embarked in the task of finding a complete mathematical and logical framework of basic concepts as a foundation of all that can be thought and said about our universe. The *Principia Mathematica*, by Bertrand Russell and Alfred North Whitehead, was part of this endeavor, as well as some of the problems posed in 1900 by German mathematician David Hilbert in Paris, at the International Congress for Mathematicians, which was mainly concerned with a complete logical formalization of mathematics in a coherent set of axioms.

Unfortunately, Kurt Gödel proved, in 1931, in what was going to be known as his "incompleteness theorem", that no complete and, at the same time, consistent set of axioms can ever exist. With this proof, the tidy Aristotelian idea of rational knowledge as an ordered and complete rendition of axioms grounding theorems and so on also came to a halt.

The next best option, the Kantian stipulation, by which an agent would consider rational only what might be acknowledged as such by a universal audience, introduces, however, a circular argument at the basis of the concept of rationality: rational is what others would accept or assent to, which presupposes my willingness to remove from the agreement everything they *would not* accept.

An important attempt in the 20th century, but not the only one, to avoid circularity in rational theory was Game and Decision Theory. Here too, however, circularity could not be entirely erased. Why? Game theory showed that in games where it is not possible to *win* the opponent, because mixed strategies and equilibrium points are valid formalizations for game theory mostly when $n=2$, collusions mediated communicatively between players *needed* to be formed.

To *cooperate* means to coordinate a common strategy that advances the self-interests of participants who have confronted positions. Already in the early days of game and decision theory it seemed pretty obvious that such cooperation could only be achieved by an explicit *communicative* rational agreement.

In the 19th century, George Sanders Peirce had also explored a different angle to understand our agreement with others, framing it in mathematical or probabilistic terms, rather than to appeal to the tools offered by traditional formal logic to force a conclusion or a valid inference on other rational agents. For Peirce, the crucial difference with the traditional philosophical way to apply general categories to particular instances was that we must be able to rigorously quantify the probability of making a sound inference.[3]

Peirce's and other probabilistic approaches to understand the acquisition of knowledge and to model our agreements with others had also very syste-

3 Cf. Peirce, Ch. S., *Studies in Logic*, Amsterdam/Philadelphia: John Benjamins Pu blishing Company, 1983 (Orig. 1883).

matic advocates in, for example, Keith Lehrer's and Carl Wagner's *Rational Consensus in Science and Society*.[4] Another example is represented by the vast and interesting amount of more recent work by Franz Dietrich and Christian List that concerns the merging of opinions and opinion pooling. Their work deals with the possibility to aggregate the individual probability assignments to some set of non-disjunctive events preserving collective probability assignment and probabilistic coherence.[5] However, all these attempts to frame analytically the possibility of rational consensuses cannot avoid the Kantian circularity either, since at least the willingness to agree with others, and to keep in place previous agreements, have to be presupposed by all the implied in the communicative process.

Habermas' concept of truth[6] defined it also as the result of a rational consensus seen as a natural consequence of conceiving the search for truth as a cooperative enterprise mediated by communication. Communication coordinates an interchange of unique visions of the world by different actors until all of them arrive to a common vision that resists persistent critical scrutiny. And this we call *truth*.

Mercier and Sperber have also defended an "interactionist" account of human reasoning in their *The Enigma of Reason*.[7] There they review experimental evidence in support to the claim that humans solve problems and optimize their individual intelligence within dialogical and argumentative contexts, without excluding the possibility of a silent scenario that can be defined as a "rehearsal" of arguments and justifications made with absent interlocutors.[8] This interactionist vision of reason is also Popperian at its core, in my opinion.

---

4 Cf. Lehrer, K., & Wagner, C., *Rational Consensus in Science and Society*. D. Reidel Pu blishing Company, 1981

5 Cf. Dietrich, F., & List, C., "Probabilistic opinion pooling generalized Part one: General agendas", *Social Choice and Welfare*, *48* (4), pp. 747-786, 2017.

6 Cf. Habermas, J., *Wahrheit und Rechtfertigung*. Frankfurt am Main: Suhrkamp Verlag, 1999.

7 Cf. Mercier, H., & Sperber, D., *The Enigma of Reason*. Harvard University Press, 2017.

8 Cf. Verssière, S., Constant, A., Ramstead, M., Friston, K., & Kirkmayer, L., "Thin king through other minds: A variational approach to cognition and culture", *Beha vioral and Brain Sciences* , *43* (90), 1-75, 2020.

2. The problem with the current boom of online social media.

Until now, and from ancient times, academia in general, and universities in particular, offered a framework of rules for supporting argumentative processes and the agreements that ensued from them. With the Internet a new deliberative agora has emerged and with it the realization that it might not be promoting exactly the cooperative search for truth.

The term "post-truth" applied to Internet and social media has been coined to designate the possibility that not every meaningful agreement can be labeled as valid or convincing. "Echo chambers" or closed social media communities that reinforce previous agreements, keeping at bay refreshing and contrasting opinions, put in evidence the perplexing fragility of the validity of our consensuses.[9]

Now we see huge numbers of educated individuals trying to form a correct picture of the world *not* arriving at the truth, insofar as they might be trapped inside an echo chamber, or a filter bubble, where they *never* are confronted with crucial contrarian evidence.

The cognitive distortions that massive social media elicit due to the sheer complexity of the conversations taking place simultaneously pose enormous challenges to the constitution of a political common will, a nonnegotiable condition for deliberative democracies.

3. Rational vs. spontaneous consensus.

Why complexity, and the spontaneous consensuses emerging from it, might distort or affect good individual and social deliberation?

As strange as it may sound, a rational consensus might be understood as a form of spontaneous consensus as well, since in principle no centralizing authority should be governing it. The theory of social contract (from Hobbes to Rawls) contemplates indeed, at the onset of the agreement leading to the constitution of the contract, the spontaneous acquiescence of all implied individuals to the necessity to subordinate one's particular political will to a leading authority, a rational "sovereign", which ought to express a common political will.

9 Cf. Manjoo, F., *True Enough. Learning to Live in a Post-Fact Society,* John Wiley and Sons, 2008, and Pariser, E., *The Filter Bubble. What the Internet is Higing from You*, New York: The Penguin Press, 2011.

"Spontaneous consensus" in applied computer science, in contrast to a rational-motivated consensus as it is studied by philosophers, has been seen in the relevant literature as the emergence of new "configurations" of large collections of elements (mostly nodes or individuals in massive social networks, but also particles if we see the emergence of the new configuration from the point of view of a physicist) that take on an "ordered form" or pattern suggesting that all particles, nodes or collection of elements, might share the same state, or might be "in agreement", or might be or "signify" the "same" thing, or might share a semantic consensus around a given concept or notion.

The main theoretical paradigm with a suitable framework for interpreting and understanding consensus in large collections of elements, be they particles, animal species, computer nodes or individuals in a social network, has been mostly thermodynamic statistical mechanics and, in particular, it has been constructed by analogy to condensed matter physics and evolutionary game theory. Evolutionary game theory, thermodynamic statistical mechanics, graph theory and condensed matter physics, applied to social networks, offer a series of conceptual strategies and tools capable to visualize and make sense of large collections of particles from a structural or topological point of view.

The idea is to be able to tackle analytically, when possible, and/or statistically, the combinatory nightmare of trying to define a mathematical function capable to deal with a physical system or an object in the world, understood as an array or a pattern of particles, using developments in particle physics that have successfully applied several mathematical tools, many of them of statistical nature, on real physical phenomena that might, by way of an analogy, be compared to clusters of individuals in agreement.

However, one of the key problems faced by models resting on statistical tools is that individuals are not simple singular entities at all, as if they were single particles, but are in themselves the result of many complex processes (physiological and psychological), which are still not well understood and that in themselves, to an extent also open to philosophical debate, are affected by particle physics. In effect, to give just one example, the dynamics of single individuals' neurological processes are still unknown, a matter of philosophical debate in philosophy of mind.

But letting this aside for the moment, the fact is that in complex systems large-scale phenomena do not necessarily depend on the microscopic details

of the process, and are not arranged or determined by the latter. They can be seen also as an independent, causally efficient, development altogether.

A good place to start exploring social dynamics as an expression of complexity is the seminal paper "More is Different", authored in 1972 by the Nobel laureate in physics Philip Anderson. As its title suggests, Anderson developed there an argument to show that truly new ordered patterns emerge when we abandon the field of structured microscopic phenomena, where we can define a molecule as the composite of a few number of particles organized in fixed or symmetric ways and polarities, and enter a macroscopic realm consisting of condensed matter seen as large collections of molecules and particles that can, as patterned conglomerates, behave in unexpectedly new ways that seem to depend on the form or pattern taken. Anderson's influential paper suggests the idea that phase transitions in large groups of particles give rise to structures truly different in themselves, which means that they are *causally efficacious* in a way not envisaged by the previous state. These structures truly become *something else*. For this reason, this paper has been considered a seminal work on complexity. The new properties emerge, according to Anderson, as the outcome of "broken symmetries".

In elementary particle physics symmetry groups can be used to classify crystal structures and other molecular transformations. It is said that a thing is symmetrical if there is something that you can do to it that let it be the same as it has been before. This means that a thing is symmetrical under a certain transformation. The first "things" that need to be symmetrical, as Einstein knew, are physical laws: they must be "translatable".

Thus for Anderson, in complex systems, the "destruction" of fundamental symmetries truly gives way to new intriguing and emergent phenomena, new symmetries that can be considered perhaps new objects. "At some point we have to stop talking about decreasing symmetry and start calling it increasing complication" he writes.[10]

As a result, we could envisage emergent clusters of elements in consensus influencing others not because there is something persuasive as such in the semantic agreements they share on social media, but because, for example and as an effect of complexity, a monster group of users clustered around a

10 Anderson, P., "More is Different", *Science, 177* (4047), 1972, p. 396.

network "hub" drags the isolated or weak-linked ones to their huge sphere of influence.

4. The analysis of equilibria in complex social systems.

One strategy to analyze consensuses in complex conglomerates of elements (such as individuals communicating in massive social networks) is to define consensus or agreements as a form of equilibrium, which allows social scientists to apply thermodynamic statistical mechanics, condensed matter physics and evolutionary game theory to the sphere of understanding spontaneous consensus.

In all sciences, including social sciences but also in philosophy (think of the Rawlsian notion of "reflective equilibrium"), equilibrium assumptions are the rule, i.e., the idea that all systems will get eventually to equilibrium and will remain in that state if suitable conditions are met. Equilibrium can be characterized in a seemingly straightforward manner (analyzing the spectral form of a matrix for example) and frequently operates as a normative or as an asymptotical approach. Equilibrium arguments are intended to offer a solution for competitive forces and strategies. But they will not always do in dynamical systems, quite on the contrary.

This is a problem already observed and tackled in political philosophy, namely, the fragility of political consensuses around a common political will. A nightmare for all those who want to design public policies in view of the future, disruption of equilibria around a political consensus can trash the best well-meaning political alliances.

In online social networks, disruptions of equilibria have been assessed with the tools offered by statistical mechanics. For example, in classical thermodynamics, any system can remain for a long time in a metastable or quasi-equilibrium state and its disruption and progression to a phase transition can be seen as a model of all other disruptions in dynamic systems. Physicist and philosopher Carlo Rovelli gives us the illustrative example of a pile of wood in a room full of air, which can be thermodynamically quite stable, remaining years just there, until perhaps a pyromaniac ignites the whole lot. The violence of the burst will dramatically increase entropy in the room, in contrast to the

initial, metastable, state. Rovelli defines metastable states as just temporary obstructions in the phase state of a system.[11]

So, from the point of view of classical thermodynamics, equilibria are just metastable states that might remain inert until a perturbation increases the entropy of the system. But from a dynamical perspective, and from the theory of non-equilibrium systems, complex systems evolve in time and can cause behaviors and form patterns whose rules we may try to understand.

Firstly, two systems might involve a feedback loop, in which two dynamical systems remain coupled with one another as metastable states function as causes for events taking place in the other system, with an internal flow of energy or information oscillating back and forth. Such a dynamics will compel us to analyze the two systems as a whole. Or, secondly, we might have a feedback system where the emergent metastable state causally determines its lower level states and crucially modify them in a sort of inverse determinism. Both perspectives could offer, each one of them, also a particular view about the formation and behaviour of spontaneous consensuses in large conglomerates of individuals and nodes in online social media.

According to George Ellis, entropy is a measure of how many lower level states correspond to a specific higher-level state.[12] So, how many ways are there to define a given state? The more ways are, in principle, possible, the more entropic is the system. Claude Shannon's definition of entropy made use of this intuition in order to devise a measure of the intelligibility of a transmitted message, where the more "uncoordinated" or diverse the microstates comprising a message are, the higher the measure of entropy will be and the *less* informative it will be.[13] Now, a higher-level state subsuming or encompassing a number of microstates can be understood as a consensus. It is easy to see what would happen in a complex system with maximum entropy: a number of microstates not able to be subsumed in a higher-level state would mean the stalemate of the system. As it turns out, then, the possibility of applying thermodynamic theory to a conglomerate of elements on social

11 Cf. Rovelli, C., "Where Was Past Low-Entropy?", *Entropy,* 21 (5), 466ff, 2019.

12 Ellis, G., *How Can Physics Underlie the Mind?* Top-Down Causation in the Human Context, Springer Verlag, 2016, p. 12.

13 Cf. Shannon, C., "A Mathematical Theory of Communication", *The Bell System Technical Journal , XXVII* (3), 1948.

media lets us understand the formation and disruption of consensus in ways altogether detached from the usual assessment based on rational agreements.

If we model the dynamics of a social network according to the rules followed by statistical mechanics, as it is usually done nowadays, the topology of massive online social networks indeed presents such obstructed paths, which may happen either by accident or design. By accident, when, for instance, the algorithms enclose a cluster of users inside an echo chamber.[14] By design, when the algorithms open the whole phase space only to users who pay for publicity. In both cases, graph theory offers a series of well-known tools to help us understand how a network can be "tweaked".

From Shannon's concept of entropy, follows that the more entropic a system is, the less informative it will be, like a disassembled puzzle. Therefore, Shannon's logarithmic formula tells us the probability of "guessing" the higher state that encompass most of the microstates within a system. The higher the entropy, the more difficult this will be, since all possible encompassing higher states would be equally likely. But, between low-entropy and a maximum entropy state, there are several possible metastable states with different patterned arrangements.

If we see an online social network under this light, complexity means for such a system that in some metastable states, the probability of finding the system under one consensus rather than under another could be at least empirically defined, irrespective of any semantic rationale. A metastable state is just a form of temporal equilibrium, as we have seen. Statistical tools could also define the probability of phase transitions from one equilibrium to another, and their probable percolation's paths, as they can be determined by graph and network theory.[15]

In complex theory, different patterns may represent simply different possible consensuses or equilibria (or metastable states). If we observe them from the point of view of information theory, we might also consider them as *causally efficacious*, that is, information would be capable to input "work" into a system.

---

14 Cf. Pariser, op. cit., and Perra, N., & Rocha, L., "Modelling opinion dynamics in the age of algorithmic personalisation", retrieved from www.nature.com/scientificreports: https://doi.org/10.1038/s41598-019-43830-2, 2019.

15 Cf. Caldarelli, G., *Scale-free networks: complex webs in nature and technology*. Oxford: Oxford University Press, 2007. Also: Newman, M., *Networks. An Introduction*. Oxford University Press, 2010.

This is the idea presented by George Ellis: complex structures are causally efficient in ways not envisioned in microscopic expressions of a physical law. Although it is true that, in the hierarchy of complexity, physics underlie all natural causes, the important question for him is not scale, but causation, and, in particular, the question about which would be the highest level of causation.[16]

That being the case, this means that context introduced by complexity *matters*. Top down causation sets *constraints* for lower level causation. It gives an order to lower level interactions and creates new possibilities whereby there was any, *restricting* others. In complex systems, the macro structures tell the microstructure how to operate.

We may also say that emergent properties, as Anderson envisaged, result in feedback loops that install a dynamism not intended by the individual components of the systems. Now the goal is created by the *information flowing* in the system.

Furthermore, quantum uncertainty says it is not possible to predict how will a system evolve, not even with all the computer power we want. These insights are crucial to understand the resulting behavior of dynamic and non-equilibrium systems and, hence, to define complexity.

5. Some empirical attributes of online social networks.

How have thermodynamic and condensed matter statistical mechanics helped us model complex collective phenomena?

Most of the first studies referred in the literature departed from the Ising paradigm for the physical understanding of ferromagnetism, one of the first attempts to model particle dynamics from a perspective informed by statistical physics.[17] It can and has been used as a simple model for opinion dynamics, where agents are seen as influenced by the state of the majority of the interacting individuals.[18]

16 Cf. Ellis, G. (28 September 2012), "On the Nature of Causality in Complex Sys tems", Lecture at the Copernicus Center for Interdisciplinary Studies. Available at https://www.youtube.com/watch?v=nEhTkF3eG8Q.

17 Cf. Castellano, C., Fortunato, S., & Loreto, V., "Statistical physics of social dyna mics", Rev. *Mod. Phys. (81)*, 591–646, 2009.

18 Cf. Baronchelli, A., "The emergence of consensus: a primer", 2018, retrieved from rsos.royalsocietypublishing.org: http://rsos.royalsocietypublishing.org/

A much used and popular method in computer simulation of social systems is the agent-based modeling. In the agent-based models, populations can be *homogeneous* (that is, the agents are identical or susceptible to be influenced in identical manner) or *heterogeneous*.

In current social sciences, where the conglomerates of interacting agents tend to be finite, the tools most used to study, from an empirical perspective, the theories and insights we can examine to define the attainment of a consensus are graph and network theory.

Network theory comes in handy due to its capacity to represent, in a graphical and straightforward manner, consensuses among agents, as it depicts them as clusters or agglomerations of agents or nodes. One of the first ways to tackle these questions on consensus referred to classical studies about "percolation", the study of random graph ensembles percolating to, or producing, new topological designs, amalgamations or configurations of individual elements.

A key structural, or topologic, feature that can be studied in the context of social networks to define probable consensuses is community detection. It pertains the task of finding groups within a network that are or may be structurally related. Seen from a purely structural perspective, a consensus can be defined as a detected community of individuals sharing a same state or agglomerated around the same cluster. A number of sociological assumptions, such as that similar individuals will tend to form part of the same group, underlie the importance of defining topologically communities within a network.

Another key problem to be tackled while studying social networks concerns the analysis of the dynamics that define the interactions among the members of a network. This type of analysis focuses on the paths taken by the information in its processes of dissemination.

The tools just mentioned can help us also define some crucial measures to evaluate consensuses in complex systems. The first one to consider is that of *centrality*. It tells us which are the truly important vertices and edges of the system. Another important mathematical measure for vertex centrality is called the *degree* of a vertex, the number of links departing or pointing to a given node.

But *the* really important measure in complex social networks is *betweenness* centrality, a number essential to understand the true influence of complex online social networks in our current world.

In a graph, the person with the bigger betweenness centrality might be also more important than the person with the greater degree. The highest betweenness centrality determines sometimes also the highest closeness centrality. Granovetter, in his seminal work "The Strength of Weak Ties", was the first scholar to point out the importance of this measure to adequately understand the structure of a social network.[19]

In effect, only weak ties can be bridges between clusters of nodes and cliques. A bridge represents the node with the largest betweenness centrality. Prior to the advent of massive online social networks, in large real social networks was rare to find such a solitary node having this role; now, if you are a social media user, they can be found everywhere. To remove a weak tie or a bridge can be very damaging to the diffusion of information, because more people can be reached through weak ties. Therefore, those to whom we are weakly tied will move naturally in other circles and have thus access to different information than those we have in our close circle of friends. The business format of online social corporations is based precisely upon their capacity to control the diffusion of information by cutting or enabling the weak ties of the whole network. This is how charging for publicity actually looks like within an online social network.

Another interesting and important attribute is that most real social networks (and all complex networks in general!) are found to exhibit a small number of "hubs", namely, few or even a couple of vertices with a unusual high degree, coexisting alongside "lonelier" nodes, that is, nodes with lesser linkages to other ones. A major topic of research in recent years has been the investigation of the effect of hubs on the performance and behavior of networked systems. Hubs formation is the effect of what it is known as "free-scale" networks, a characteristic probability distribution in complex systems, where the number of edges connecting the vertices cannot be distributed around a mean and whose degree distribution falls to zero slowly, in a long-tail or power law distribution.[20]

A further important recent discovery in the study of opinion dynamics is the effect produced by what are known as "zealots".

---

19 Cf. Granovetter, M., "The Strength of Weak Ties", *American Journal of Sociology* ,*78* (6), 1360-1380, 1973.

20 Cf. Caldarelli, op. cit.

Zealots are individual nodes that never change their minds: biased individuals who favor unrelentingly one opinion.[21] In empirical scenarios, they might be represented by users of social networks who, no matter what kind of evidence you might present to them, will never change their minds, remaining always adamant to their own prejudiced opinions.

Many studies have found that in the presence of zealots branding just one opinion, zealot nodes would quickly dominate. A few zealots would end up swaying the whole outcome. If, on the other hand, they had equal fractions of zealots for all kinds, that is, every opinion had people who adamantly believed in that opinion, the model reached quickly a stalemate or a high entropic state. According to Raissa D'Souza, the more opinions you have that are defended by zealots, the more we can find fractions of the population that will never reach consensus, but stale (*and* maybe polarize?).[22]

On the other hand, when you have many different opinions and zealots for all of them, it is possible to actively fracture a political status quo or equilibrium. And here, according to D'Souza, even when it is indeed possible to overthrown a ruling party, the studies show that it seems impossible to reach a consensus again: a group of political minorities remain stuck and forced to coexist in an unstable non-equilibrium state. Then a divided opposition might fail to create a stable future government. This has been observed many times in history, adds D'Souza.[23]

6. The modeling of complex online social systems and its consequences for the diffusion of truth statements in current societies.

The studies I have referred so far suggest that universal consensuses in complex systems of opinion dynamics that can be objectively analyzed as online social networks, where information circulates, depend on the paths taken by the flow of information within networks' topologies, whose structural aspects must result from constraints of *physical* nature, and are, therefore, not subjected to rational, cognitive or institutional control. Identifying and understanding the physical factors that may affect or obstruct the flow of

21 Cf. Mobilia, M. "Does a single zealot affect an infinite group of voters?", *Phys. Rev. Lett.* (91), 2003.

22 Cf. D'Souza, R., "The Mathematics of Influence and Opinion Analysis". Lectu re at the UC Davies Social Sciences, 2015. Available at: https://www.youtube.com/watch?v=VKGsRU5yBzc.

23 Cf. D'Souza, op. cit.

information in these systems is crucial for the sustainability of democratic institutions defined by their deliberative character, specially there where people effectively use online agoras to deliberate alongside others or to exercise their rights as reflective citizens.

The same kind of difficulties could be said that emerge from a lack of agreement among economic agents with respect as to how to interpret the information they receive. This is the subject of an article by the former Finance Minister of Argentina, Martín Guzmán, and Nobel Laureate (and his teacher) Joseph E. Stiglizt: Guzman and Stiglitz: "Towards a Dynamic Disequilibrium Theory with Randomness".[24]

The article is interesting because the authors understand market equilibria *as a consistency of the beliefs of the economic agents around macroeconomics policies and perspectives*. To attain this sort of equilibrium of beliefs, or to keep a macroeconomic consistency, requires, according to Guzmán and Stiglitz, both objective and, of course, subjective conditions (or similar priors), that is, an already assumed or previous measure that some events will occur or not with a given probability.

The authors stress out that it is the aggregated harmonization of perceptions and beliefs of multiple agents, whose plans are interconnected in an economic space, which could give ultimately consistency to the effective coordination of their reciprocal plans and projects for the future. This is the crucial subjective condition whose absence could maim the capacity of an agent to access to credit, and that can provoke a crisis. Changes in beliefs can be produced by many factors. Any new circumstances can in fact dramatically alter the posterior probability distribution of any future state for any participant in the market. This implies that many changes in beliefs do not necessarily depend on real events, but can have however very real effects.[25]

Yet the a priori presupposition of stable universalistic agreements among social agents, including the validation processes affecting the legitimacy of our social institutions, assume non-evolving equilibrium systems for the attainment of consensus, where "false" or rejected conceptions are, so to say, cancelled out by the "true" ones in the space of different theories about the world (including a social and subjective world), while remaining nevertheless

24 Cf. Guzman, M., & Stiglitz, J. E., "Towards a Dynamic Disequilibrium Theory with Randomness", *NBER Working Paper* (27453), 2020, https://www.nber.org papers/w27453.

25 Guzman & Stiglitz, op. cit., p. 22.

available when necessary, and where feedback loops between false and true, or more probable and less probable, hypotheses are not possible. But in evolving non-equilibrium systems, which are multidimensional, it can happen that theories or opinions rejected by a group can truly disappear, or even be erased within the manifold, leaving behind the cost that would be entailed if they were needed to support a true or valid statement.

If we see consensuses as a physical phenomenon, it seems just natural to suppose that in a space of competing assertions about a state in the world, different possible consensuses will naturally absorb degrees of freedom while the social platforms model them analytically. As a result, several constraints will be imposed on the model of the world every agent constructs along with others when they engage in complex communicative exchanges on these social platforms.

In effect, if you have the mean of a sample, the bigger it is, the more room for error you will have. But, at the same time, the more data you have, the more pieces of independent information you will need to successfully manage them. So every time you make a calculation, like the mean of a given sample, independent pieces of information will be "used up" in order to calculate the given mean of a category, in detriment of others. The so-called curse of dimensionality[26] can be also associated to the kind of phenomena arising when you have too many categories or dimensions to deal with, when defining a function.

The exhaustion of degrees of freedom, while making a model for our data within online social platforms, will impose inevitably *constraints* to the information available at a given moment. It seems that having to "crunch" dimensions in a low dimensional model to correct error measures within a space of online opinions, not only leave aside information to which we attribute, maybe, a low probability; the information that does reach our probability threshold will inevitably create a phase space where the rejected information will be constrained. Not in mathematical, but in physical terms, the analogy can be extended further to suggest, as George Ellis does, that in the brain, which is also a complex information-processing system, signaling molecules alter constraints that shape the flow of information, transforming them dynamically. What happen with the previous configurations that were "discarded" over time? They are "forgotten", so to say, "erased" or left behind: "…all the

26 Cf. Bellman, R., "Dynamic Programming", *Science, 153* (3731), 1966.

molecules argy-bargy every microsecond means that the details of their initial state of motion are irretrievably lost".[27] Although it is true that higher-level processes depend on the lower-level physical processes that underlie them, this does not mean that the lower-level processes govern the emergent ones in deterministic fashion.

The erasing of relevant information when a group of people attempts to get to a common vision of the world seems to point out to the fact that truth, being the result of a rational intensive process, could be attained, or assented to, always within a restricted group of individuals, as it will be seen below. Historically we have been already in this situation. We have known for centuries that the earth was not flat, having Eratosthenes calculated its diameter with pretty accuracy around the II century AC. However, a majority of people still believed throughout the centuries that the earth was flat, a majority that coexisted alongside a minority of privileged medieval scholars that knew better simply because they had access to *restricted* information. Thus, the problem we have cannot be simply solved by implementing a syllabus of universal knowledge across our school system, although the homogeneity of learning spaces could be of help, as we will also see. My problem pertains not simply psychological, empirical, historical or cognitive limitations, although evidently these also play an important role. It is a problem that can be formulated as the impossibility of sharing attributes over the world, among these the attribute of being true or valid, and attaining consensuses, in a truly universal sense or with universal scope, in our physical world when complexity as we know it reaches some threshold.

7. Problems posed by "curse of dimensionality", the bias-variance tradeoff and non-equilibrium systems.

Let us see this in more detail.

In mathematical terms, to say that two or more agents are in "agreement" or in "consensus" means that there is a function $f(x)$ that represents the agents, seen as data points on a geometrical space, as sharing the same "state" (this usually means they share some traits or attributes in common).

Therefore the problem here is, above all, to define *how* we are going to *represent* our data when modeling a space of shared opinions. Otherwise, we would be forced to ask a mathematical function (a "machine") to find traits in

27 Ellis, G., "From Chaos to Free Will", *Aeon* Magazine, June 9, 2020. Available at: https://aeon.co/essays/heres-why-so-many-physicists-are-wrong-about-free-will

common between any two data points and learn whatever it could, or to find any possible clusters of data points with similar attributes. But such model-free or tabula rasa approach would only succeed if one had unlimited time and data samples, if this is possible at all.

This task will be inevitably confronted with "the curse of dimensionality", an anomaly that emerges when data is organized and analyzed in multidimensional geometrical spaces (a.k.a. hyperplanes). It says that the more dimensions (or attributes) a model of a space of opinions has, the more data we will need to "train" it, until the space of points grows so large that the whole thing becomes unmanageable.

Rigorously speaking, it becomes impossible to calculate the mean square error –a reliable measure of error for mathematical functions– of any data point, since the proportional difference between the farthest-point distance and the closest-point distance, which would give the necessary contrast to estimate the precision of the function, "vanishes".[28] Within an algorithm attempting to define clusters by nearest or neighboring data points expressing opinions, this implies that distances to the mean cease to be usable as a measure to discern or discriminate between near and far neighbors.

If we had a multidimensional dataset, the curse of dimensionality means that to define the error of the function with respect to the precision of our attribute classification we would need to compute almost the whole dataset to avoid bias in our final classification *for each dimension*, which amounts to say we would not have a mean distribution of the values for *any* dimension, for the reason already mentioned: the absorption of degrees of freedom.

Although many clever algorithms have been created to circumvent the challenges arising from dimensionality, or to reduce it into a space of low dimensionality, high dimensionality expresses something about our physical world that defies all attempts to define a space in which, if you allow me to put it this way, we all can be met around the same mean.

This is not much different from what really happens when someone wants to understand or reach an agreement with other people. When attempting to communicate even highly formalized scientific papers (which presuppose a standardized language that everybody should understand beforehand), the author or authors cannot avoid trying to convey their contents appealing to some already shared conceptual context, without which it would be very difficult to make themselves truly understood. This is why scientific discove-

28 Zimek, A., Schubert, E. and Kriegel, H. P., "A Survey on Unsupervised Outlier Detection in High-Dimensional Numerical Data", *Statistical Analysis and Data Mining*, 5, 363-387, 2012, p.2 https://doi.org/10.1002/sam.11161.

ries normally require a lengthily assimilation period, where scientists learn the common semantics of the relevant concepts that define a new theory, before accepting it. "Negotiating" a basic conceptual framework that offers a context for an agreement is an unavoidable condition for every rational discourse and this is why we say truth is the result of an intensive process.

However, on the other hand, computer science teaches us that a consensus or an understanding cannot be *too good* either. If an agreement is very "successful" it might be that such consensus will not be able to be extended outside a reduced group of people very familiarized with all the nuts and bolts of a theory on which all of them have agreed. They might understand each other or be "attuned" to each other simply *too* well. Here the theoretical explanation will not be "generalizable" enough, running the risk of remaining merely as an esoteric one.

A good algorithm should be unbiased, since it cannot define a priori a particular solution. However if you don't have bias, you'll have high variance. A function flexible enough to approximate a wide range of input/output mappings will be by definition very sensitive to the idiosyncrasies of the particular dataset used for its training, and can therefore run the risk of "overfitting" the data, that is to say, given enough time, it will be able to learn all the characteristics of its particular training set until it covers them completely, making it unable to generalize to non-previously observed cases.

The need to balance bias and variance is an unavoidable challenge when defining a function that classifies agents as belonging to the same group and/or to express an agreement. Both of them "adapt" the function in two different "too good" ways to the dataset: either by choosing an attribute of the dataset in detriment of another important or relevant one, or by mapping all the attributes of a sample in detriment of the possibility of generalizing the function.

The bias-variance tradeoff is not simply a mathematical curiosity: it seems that indeed the tabula-rasa model of knowledge, where no bias is introduced, simply does not exist in the real world.[29] When communicating with others we do need some bias, that is, some context for the mutual understanding. This is known since antiquity, where Aristotle represents the first syste-

29 Cf. Geman, S., Bienenstock, E. and Doursat, R., "Neural Networks and the Bias/Variance Dilemma", *Neural Computation*, 4, 1-58, 1992. Available at: https://www.ics.uci.edu/~smyth/courses/cs274/readings/bias_variance.pdf.

matic attempt to define our biases as categories comprising our ontological assumptions of the world. The problem we face now in complex societies is that each semantic context is per definition too restrictive. On the other hand, online social platforms sometimes, when acting in good faith (which, of course, might be not always the case) may have no other choice than to choose for us the biases with which we will understand each other.

That being the case, the impossibility to reach an agreement without introducing some bias, which amounts to being obligated to choose some traits over others, some attributes over others, based on domain knowledge or plain and simple prejudice, seems unavoidable. As it seems unavoidable too the conclusion of the sheer impossibility to expect that the same biases will function as catalyzers for consensus with every kind of people. That is to say: some bias will be needed to be understood by a group, some other to reach an agreement with another. This will put in place some unavoidable *constraints*. This seems also to suggest that, in effect, in the face of real complexity, like the ones human groups are beginning to experience in their day-to-day interactions, cognitive and practical agreements could become less and less generalizable.

8. Truth, meaning and the flow of information.

A semantic meaning specifies the meanings of the words, what is the specific meaning of a word for a person or a group. A related question asks in virtue of which attributes or characteristics of a person or a group, and the world they inhabit, a word has a specific meaning for that group. In general, a semantic theory is the description of meanings of words within a given symbolic system. We may also define it as an ontological theory of the world. And here we can observe that there seems to be a natural language semantics that is common to many different languages and that may be rooted in a common ontology of the world, which philosophers throughout the centuries have defined as categories or basic concepts.

The ontological aspects of semantic theories have been addressed by theories of semantics that are not merely descriptive of the empirical concepts found within a given culture. A sentence will have a meaning, or not, depending on whether it is true or false, and this depends, in turn, on the accuracy of the information about the world a given sentence encodes or

expresses. Another way of stressing this point is to say that sentences define always the conditions and rules that make them true or false.

A classical theory of meaning in this sense, by its own definition, seems to exclude a priori a truly universal agreement in relationship to all the conditions and rules that make sentences true, save perhaps in a limited number of assertions about the world, such as, for instance, "now the sun is rising". But many other assertions are simply impossible to universalize without sharing with your interlocutors a common world.

But it is true that from the fact it is very difficult to provoke an assent around a given true sentence amidst interlocutors who do not share with you a common understanding of the world, does not follow that the sentence might be false. Another way of conveying this idea would be: a true assertion can *always* be *in principle* universalizable. Perhaps it is very difficult to convince a group of enthusiastic earth-flat supporters that the world is a globe, but this difficulty does not make the latter less susceptible to be universally shared. And you could add, as Kant did: a group of earth-flat supporters, if they are rational, will eventually come to understand that the assertion "the earth is round" is susceptible to be *universalized* in a way the other cannot possibly be.

The assumption that truth can be shared by a universal audience became a decisive mark for truth in modern philosophy ever since Kant formulated his categorical imperative, which asserts that something, to be rationally acceptable, must be universalizable. We have already noted the circularity inside this definition, since it asserts that something, to be taken as true or valid, must be considered true or valid for every rational being pondering dispassionately the matter. But if our intuitions about the truth of an assertion are sound, in the sense that a group of speakers must share more or less the same information about the world in order to concur around it, it would physically be impossible to ever be able to find an audience who could share with you all the relevant information that would make a statement persuasive enough to be considered true. Therefore, we may ask: what is the real import of a universalistic stipulation for truth?

A Kantian philosopher could retort that universalization is a regulative idea, or an asymptotic condition: if all the relevant objective conditions were met, then everybody could see that a given assertion is true; in the same way it is easy to understand that even if the sun is not rising now, it might very well be rising in China.

But universality as a relevant mark for truth in modern philosophy is not simply a regulative idea, or an aspiration. An assertion whose truth cannot be assessed by at least a majority of rational beings will always remain suspicious. A scientific result that cannot be replicated by other laboratories will be discarded. A Miss Universe that hasn't competed with contenders coming from all over the world will be a scam. Truth has always been considered susceptible to be shared by the majority of people, if not by *all*. As P.C.W. Davies has aptly suggested, we have a "Platonic" conception of truth, as physical laws also must be, that is, infinitely precise, perfect, immutable, and defined by mathematical relationships that transcend and describe the universe regardless of its empirical or contingent circumstances.[30]

From the behavior of online social media users seems to follow, however, that susceptibility to be shared by a whole community of argumentation (any possible rational being or an impartial rational being) cannot be considered a condition for the truth or validity of a statement or judgment about the world, including our intersubjective or normative social world. And this not only because there are epistemic conditions that are demanded for a rational assessment to ensue (such as, as already mentioned, a common experience of the world or a shared semantics), but also because the same physical information cannot be disseminated evenly or homogeneously in complex communicative contexts.

Thus, if meaning and truth, or semantics in general, are encoded in information systems, the conclusion ought to be that information limits necessarily the scope of meaning and truth. And this, in my opinion, is a serious objection to our current understanding of truth. Because we tend to suppose that truth is more "rationally shareable" than falsehoods.

Suppose someone says "if we build a ship that surpass the velocity of light we would be able to arrive at Proxima Centauri (the nearest star to our solar system, which is 5 light-years away from us) in less that five years and maybe come back in time for the next Football World Cup". Such a sentence does not have really true meaning, simply because there are not many communities of argumentation in which this statement would stick as an interesting one, unless it is used for a science fiction plot.

30 Cf. Davies, P.C.W, "The implications of a cosmological information bound for complexity, quantum information and the nature of a physical law", https://arxiv.org/abs/quant-ph/0703041, 2007.

However, if there are not many, it is an empirical fact that there could be some. If they want to grow their number of affiliations, they would have to construct a consensus around their vision of reality with the potential new members. Information, in contrast to Platonic truths, can be bended in order to sustain the common vision of the world a group has come to share. And the behavior of massive online social networks, as we have examined them here, suggest that there is always a threshold, empirically definable, over which the sharing of a common vision of the world, necessary to sustain consensuses around the meaning of sentences, comes to a halt. This suggests that rational assent has an individual cognitive component that requires the exercising of a personal and autonomous reflection. This personal component, which is also a first-person assessment on what is being rationally discussed, is the core of many semantic theories that do not rely on universalization for the foundation of statements. In a given space of opinions some statements are more "shareable", easier to share, than others, but this flow can be as smooth or as rough as the extension of the shared system of basic system of beliefs about the world permits, and will depend on to what extent a space of opinion exhibits homogeneous subjects.

Therefore, susceptibility to be universalized, or the Kantian rational condition, is contingent to the truth or validity of a statement for a group of people that share the same information about the world. To see this, let us now consider the following fact, which I would call the "Lucas Conjecture" (from George Lucas, the movie director and producer).

When the first film of the saga Star Wars was released in 1977, what impacted its spectators the most all over the world was the realistic depiction of the extraterrestrial environment and spaceships, its "special effects". I personally remember how thrilled my two younger brothers were about them. They and their friends were eager to watch the movie again and again because of the exact and rigorous representation of what would be a war between two interstellar powers, which never ceased to amaze them.

George Lucas and his collaborators understood that, in fictional movies, in order to lure the spectators into the credibility of the dramatic plot, and thus awaking their emotional responses, it is necessary to offer the most realistic depictions of the environment in which the plot develops. That is to say, it is necessary to mimic the context that sustains the authority of what is being said. This happens for the reasons highlighted in this paper: because in order

to be able to say that a statement about the world is true, you had to have agreed with a number of epistemic preconditions, among these the eventual assent to what others in a community of argumentation believe. In the film, you have a director and a team asking you to really believe that a war between extraterrestrial powers is taking place, so you can resonate with what they are trying to convey to you through their plot.

Now imagine it is the BBC, a media outlet that some people consider 100% reliable and not biased, that is announcing, at the 9 o'clock news segment, that an alien spaceship has been glimpsed somewhere at the outskirt of the solar system. Because of its a priori credibility, a blurred image of something vaguely resembling an alien spacecraft would do. Even the tiniest cast of a dark shadow on the shape of Neptune will do to awaken a worldwide excitement akin to what my brothers and their friends felt when they watched Star Wars for the first time. This is precisely why consensuses and agreements are so important for truth, since they offer the context of semantic meanings in which a community can consider a statement or a judgment true or valid. A credible BBC distinguishes itself from a science fiction film insofar as we suppose that the BBC is in agreement with rigorous scientific knowledge on what it is happening right now in the world, whereas the team producing Star Wars was attempting to film a science fiction. Lucas and his collaborators had to work harder in order to gain our trust. This is what I have called the Lucas Conjecture: the importance of consensus of information and previous agreements about the world for the establishing of truth or, in this case, truthfulness.

Both contents, the blurred spacecraft and the spectacular special effect, are an ordered pattern of information subjected to thermodynamic principles. But one is undoubtedly more robust than the other. Strangely enough, the most robust pieces of information are not the ones that appeal in the most salient way to your senses, namely, the amazing computer-generated spaceship, but the ones that enjoy already of contextual informational consensus: the blurred shape showed by the BBC at the 9 o'clock news broadcast.

This is why the current boom of social media has been so devastating for democracy in many countries. Now it is possible for an interested individual or party to pay an army of cyber troops to lure you into believing that some fake news possesses an incredible amount of support among the electorate. It is possible to use the complexity aroused by the new deliberative agora in

order to convince, and confuse, thousands or even millions of uneducated individuals into believing that some wacky theory about your world is true.

It is an empirical question to decide to what extent such efforts have been successful so far. But I can advance a personal opinion, based on my experience as a citizen of a country divided by a Twitter Wars: there is a limit to what these attempts at manipulating public opinion can achieve. The reason rests upon what I have been tackling on this paper: there is a physical threshold above which a piece of information cannot colonize the whole space of opinions. The more complex and diverse a space of opinions is, the more fragmented it will be, until it reaches, at least theoretically, maximum entropy. Of course, a partisan community manager could always say that she would be contented with a majority opinion. However, this majority opinion would always be unstable and susceptible to be disrupted by an opposing opinion, as it happens now. The result, as it is predicted by thermodynamics, is a perennial or unrelenting fragmentation and polarization, typical of dynamic systems or systems in non-equilibrium.

However, as what I have called the Lucas Conjecture suggests, the metastable state that is a consensus around a belief does not depend entirely on the dynamic flow of information, but above all on the semantic meaning that is created within a community sharing a common world encoded by linguistics conventions. More than suitable channels to share a vision of the world are needed in order to find a consensus. This surplus is what each of us has come to believe as true or meaningful in our *real* dealings with the world; that is, the personal experience that enriches our self-reflection.

This is the blind spot that cannot be always acknowledged by those who think that being followed by millions is enough to flip a space of opinions favoring a political stance.

In relation to this, Vicente-Page et al. have shown that changing dynamics in a group promotes maximal accuracy *rather at a small group size*.[31] This is a result of a dynamics where individuals re-evaluate their first, probably inadequate, solitary or individual judgment, again when they are already within the group, with more access to social information. Since this does not imply that in large groups the improvement will be necessarily higher, because errors

31 Vicente-Page, J., Pérez-Escudero, A., & Polavieja, G., "Dynamic choices are most accurate in small groups", *Theoretical Ecology*, Vol. 11, pp. 71-81, 2018. https://link.springer.com/article/10.1007/s12080-017-0349-9.

could spread also rapidly there, not giving the individual the possibility to evaluate the social information in a timely manner, the maximal result in accuracy will be reached at a *small* size of the group. Their results point out to a small or intermediate size of the group as an optimal threshold to equilibrate individual and group accuracy:

> We show that the dynamics in which individuals in groups can reconsider their choices leads naturally to the result that groups of small size choose best. Re-evaluation allows a better access to social information to all members in the group, resulting in a better decision accuracy but less so for large groups in which re-evaluation does not stop chains of errors. The highest accuracy then corresponds with groups of intermediate size, and *this optimal size decreases as the influence of the social information increases.*[32]

Thus, collective accuracy does not grow monotonically with group size. That is, although the information gathered by a group could outperform the information on which a single individual bases his decisions, beyond a certain threshold the quality of the decisions made by the members of the group will deteriorate.[33] Furthermore, if the information is too complex (not a single cue concerning a location of a source of food, for example), beyond a certain point it will be increasingly difficult for a single person to convey to others exactly what she means.

It is just natural then to suppose that, in human societies, the capacity to articulate reasons for or against a decision, and to help others come to a sound or rational decision, is restricted to a medium-size group, at most. We have examined in the previous sections why this might be so. In online social networks, where, for the reasons I have already analyzed, the platform's engineers are forced to juggle with physical constraints such as the bias-variance tradeoff and the multidimensionality of massive datasets, when modeling their algorithms a certain capacity to universalize contents and meanings will be unavoidable lost.

It seems then, that, in contrast to Kant's intuitions on the universality of good reasons, truth will inevitably be restricted to medium-sized groups, where the individual assessments of facts can be preserved and effectively

32 Vicente-Page et al., Ibid., p. 72, emphasis added.

33 Cf. Kao, A., & Couzin, I., "Decision accuracy in complex environments is often maximized by small group sizes", *Proceedings. Royal. Society*, Vol. 281, 1784, 2014. *https://royalsocietypublishing.org/doi/10.1098/rspb.2013.3305*.

transmitted. This is, of course, different from affirming that everything that a medium-sized group believes can be considered true. It suggests simply that some individuals and groups are better informed or nearer to better sources of information than other individuals and groups. The Twitter user that screams his or her unreflective convictions to anyone willing to hear would be well advised to bear this in mind. Complexity can hinder more and more the possibility not only to amplify a consensus around a true statement in universal sense, but also can trapped anyone inside collective consensuses around blatantly wrong statements. Therefore, the universality condition, understood as a mark for truth, not only is not enough to support a rational assessment: it is rather inadequate for the task of sustaining rational agreements.

To be in agreement with others depends on conditions that require the sharing of a common world that cannot be completely expressed through low dimensional matrices of bags of words, as online social media platforms are forced to do in order to model conversations. This is the lesson, I think, we could derive from the fragmentation of opinions that characterizes the online social media environment.

9. Conclusions.

In complex environments, including complex societies, global or even majority consensuses might be then unattainable, unless actively pursued, and truth, as an attribute of a set of beliefs, might be forever enclosed within the group with good access to modular and, therefore, restricted information.

Nonetheless, rationality is a demanding process. It seems that, in complex social networks, above an empirical threshold the quality of the information gathered by an individual or a group of individuals with privileged access to good information will deteriorate and, therefore, will not command our consensus or agreement. Or it will be confronted by perhaps good but also biased information gathered by a group at the other extreme of the dynamic system, or, paraphrasing P.C.W. Davies, in a "region" not connected to us. We might not even find a common soil on which we can begin to talk, because our important sensorial cues could have had simply disappeared from another person's vision of the world, constraining our capacity to make ourselves understood. As the medieval scholars who profited from the discovery of America proved, only a tiny group of people could have access to information that make all the difference between a good and a miserable life. In fact, Italian

philosopher Umberto Eco thought we were heading to a new medieval era, where the shared scientific, moral and political intuitions that created the institutional cohesion of modern societies would be irretrievably lost or enclosed within the cognitive equivalent of feudal dominions.[34]

However, some of the studies we examined seem to offer a way out of the conundrum.

We can interpret topological homogeneity in online social media as a uniform susceptibility of all agents to be persuaded. This could mean a uniform cultural and educational background, where no one is left outside a formal and basic education system. It also means an efficient way to aggregate individual preferences: a good voting system, a society that vigorously resists electoral abstention.

Another unavoidable path forward is the control and submission of online social media platforms to social scrutiny. I think online social media platforms should be integrated in the defense of the ideals of deliberative democracies.

The difficulties examined by me are truly physical problems. The situation is comparable to what happens in economy: the state has to intervene when economic anomalies and crises emerge from the uncoordinated activities of millions of economic agents, each of them harboring goals and strategies of their own. Now, as in the economic sphere, millions of individuals can articulate and broadcast their opinions in the same way they have cash in their pocket to buy groceries (or not). In contrast to what the situation was a mere 50 years ago, where the majority of people sat passively in front of a TV set, now millions of people can express what they think and find a resonating chamber somewhere in the social network. This possibility is what produces the different phenomena that complexity theory has outlined in other fields where large conglomerates of elements interact with each other.

In general, a well-articulated formal education, accessible and mostly free, as well as a lucid and knowledgeable public control over the black box that are the data managing centers of social platforms corporations (where they are at the mercy of non-elected scholars and researchers –some of them of paramount academic prestige– obeying the guidelines of their private employers), are key issues for the purpose of avoiding further fragmentation and isolation of large segments of the population in contemporary societies.

34 Cf. Eco, U., *Travels in Hyperreality. Essays*, Harvest, Harcourt, 1986.

Prof. Dr. Luz Marina Barreto Guerra
Escuela de Filosofía
Universidad Central de Venezuela
E-mail: luzmbarreto@gmail.com

References.

Adorno, T. et al., *Der Positivismusstreit in der deutschen Soziologie*, several editions, 1969.

Anderson, P., "More is Different". *Science , 177* (4047), 1972.

Baronchelli, A., "The emergence of consensus: a primer", 2018. Retrieved from rsos.royalsocietypublishing.org: http://rsos.royalsocietypublishing.org/.

Bellman, R., "Dynamic Programming". *Science, 153* (3731), 1966.

Caldarelli, G., *Scale-free networks: complex webs in nature and technology*. Oxford: Oxford University Press, 2007.

Castellano, C., Fortunato, S., & Loreto, V., "Statistical physics of social dynamics", *Rev. Mod. Phys.* (81), 591–646, 2009.

Davies, P.C.W, "The implications of a cosmological information bound for complexity, quantum information and the nature of a physical law", https://arxiv.org/abs/quant-ph/0703041, 2007.

Dietrich, F., & List, C., "Probabilistic opinion pooling generalized. Part one: General agendas", *Social Choice and Welfare , 48* (4), 747-786, 2017.

D'Souza, R., "The Mathematics of Influence and Opinion Analysis". Lecture at the UC Davies Social Sciences, 2015. Available at: https://www.youtube.com/watch?v=VKGsRU5yBzc.

Eco, U., *Travels in Hyperreality. Essays*, Harvest, Harcourt, 1986.

Ellis, G. (28 September 2012), "On the Nature of Causality in Complex Systems", Lecture at the Copernicus Center for Interdisciplinary Studies. Available at https://www.youtube.com/watch?v=nEhTkF3eG8Q.

Ellis, G., *How Can Physics Underlie the Mind? Top-Down Causation in the Human Context*, Springer Verlag, 2016.

Ellis, G., "From Chaos to Free Will", *Aeon* Magazine, June 9, 2020. Available at: https://aeon.co/essays/heres-why-so-many-physicists-are-wrong-about-free-will.

Geman, S., Bienenstock, E. and Doursat, R., "Neural Networks and the Bias/Variance Dilemma", *Neural Computation*, 4, 1-58, 1992. Available at: https://www.ics.uci.edu/~smyth/courses/cs274/readings/bias_variance.pdf.

Granovetter, M., "The Strength of Weak Ties", *American Journal of Sociology , 78* (6), 1360-1380, 1973.

Guzman, M. M., & Stiglitz, J. E., "Towards a Dynamic Disequilibrium Theory with Randomness", *NBER Working Paper* (27453), 2020. https://www.nber.org/papers/w27453.

Habermas, J., *Wahrheit und Rechtfertigung*. Frankfurt am Main: Suhrkamp Verlag, 1999.

Cf. Kao, A., & Couzin, I., "Decision accuracy in complex environments is often maximized by small group sizes", *Proceedings. Royal. Society*, Vol. 281, 1784, 2014. *https://royalsocietypublishing.org/doi/10.1098/rspb.2013.3305*.

Korsgaard, C. M., *The Sources of Normativity*. Cambridge University Press, 1996.

Lehrer, K., & Wagner, C., *Rational Consensus in Science and Society*. D. Reidel Publishing Company, 1981.

Manjoo, F., True Enough. Learning to Live in a Post-Fact Society. John Wiley and Sons, 2008.

Mercier, H., & Sperber, D., *The Enigma of Reason*. Harvard University Press, 2017.

Mobilia, M., "Does a single zealot affect an infinite group of voters?", *Phys. Rev. Lett.* (91), 2003.

Newman, M., *Networks. An Introduction*. Oxford University Press, 2010.

O'Neill, O., *Constructions of Reason. Exploration of Kant's Practical Philosophy*. Cambridge University Press, 1989.

Pariser, E., *The Filter Bubble. What the Internet is Hiding from You*. New York: The Penguin Press, 2011.

Peirce, Ch. S., *Studies in Logic*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1983 (Orig. 1883).

Perra, N., & Rocha, L., "Modelling opinion dynamics in the age of algorithmic personalisation", 2019. Retrieved from www.nature.com/scientificreports: https://doi.org/10.1038/s41598-019-43830-2

Rawls, J., *The Theory of Justice*. Harvard University Press, 1971.

Rovelli, C., "Where Was Past Low-Entropy?", *Entropy , 21* (5), 466, 2019.

Shannon, C., "A Mathematical Theory of Communication", *The Bell System Technical Journal , XXVII* (3), 1948.

Verssière, S., Constant, A., Ramstead, M., Friston, K., & Kirkmayer, L., "Thinking through other minds: A variational approach to cognition and culture", Behavioral and Brain Sciences , 43 (90), 1-75, 2020.

Vicente-Page, J., Pérez-Escudero, A., & Polavieja, G., "Dynamic choices are most accurate in small groups", *Theoretical Ecology*, Vol. 11, pp. 71-81, 2018. https://link.springer.com/article/10.1007/s12080-017-0349-9.

Zimek, A., Schubert, E. and Kriegel, H. P., "A Survey on Unsupervised Outlier Detection in High-Dimensional Numerical Data", *Statistical Analysis and Data Mining*, 5, 363-387, 2012, p. 2. https://doi.org/10.1002/sam.11161.