In: *The Logic of Social Practices. (Studies in Applied Philosophy, Epistemology and Rational Ethics 52)*; Giovagnoli, R., Lowe R., Eds.; Springer Nature: Berlin-New York, 2020, pp. 87-119

The Computational Challenge of Amartya Sen's Social Choice Theory in Formal Philosophy

Gianfranco Basti[1]✉, Antonio Capolupo[2] and Giuseppe Vitiello[2]

[1] Fac. of Philosophy, Lateran University, Piazza S. Giovanni in Laterano 4, 00184 Rome, Italy
basti@pul.va
[2] Dept. of Physics, University of Salerno, 84084 Fisciano, Italy
capolupo@sa.infn.it; vitiello@sa.infn.it

Abstract. A significant chapter of the short history of formal philosophy is related with the notion and the theory of the so-called "social welfare functions (SWFs)", as a substantial component of the "social choice theory". One of the main uses of SWFs is aimed, indeed, at representing coherent patterns (effectively, algebraic structures of relations) of individual and collective choices/preferences, with respect to a fixed ranking of alternative social/economical states. Indeed, the SWF theory is originally inspired by Samuelson's pioneering work on the foundations of mathematical economic analysis. It uses explicitly Gibbs' thermodynamics of ensembles "at equilibrium" based on statistical mechanics as the physical paradigm for the mathematical theory of economic systems. In both theories, indeed, the differences and the relationships among individuals are systematically considered as irrelevant. On the contrary, in the mathematical theory of "Social Choice Functions" (SCFs) developed by the Nobel Laureate in economy (1998) Amartya Sen, the interpersonal comparison and the real-time information exchanges among different social actors and their environments – different ethical values and constraints, included - play an essential role. This means that the inspiring physical paradigm is no longer "gas" but "fluid thermodynamics" of interacting systems passing through different "phases" of fast "dissolution/aggregation of coherent behaviors", and then staying persistently in "far from equilibrium conditions". These processes are systematically studied by the quantum field theory (QFT) of "dissipative systems", at the basis of the physics of "condensed matter", modeled by the "algebra doubling" of coalgebras. This coalgebraic modeling is highly significant for making computationally effective Sen's SCF theory, because both based on a dynamic and not statistical weighing of the variables for interacting systems, respectively in the physical and social realms.

Keywords: Formal Philosophy. Social Choice Theory. Quantum Field Theory.

1 Introduction

1.1 The Social Choice Theory: a short history

A significant part of the mathematical theory of social, political, and economical sciences, specifically in the context of welfare politics and economy, is related with the development of the notion and the theory of the so-called *social choice theory,* in which the theory of "social welfare functions (SWFs)" occupies historically a primary place.

Social choice theory, as far as a formal theory, constitutes today a branch of the so-called *formal philosophy.*

This is a newborn discipline that, despite the name, started within the theoretical framework of computer sciences, essentially for the stimulus of the AI research, for migrating during these very last years from the computer science departments into the philosophy departments too (see [1] for an updated introduction). Substantially, it consists in the formalization of philosophical theories using the axiomatic method applied not only to *mathematical logic* – as it was since the beginning of the second half of XIX cent. – but also, to *philosophical logic.* Indeed, different philosophical theories (ontological, epistemological, ethical, social) consist ultimately into different *models* or *semantics* of the axiomatic *modal logical calculus,* (that is an extension of the mathematical logical calculus), and then into different *interpretations* of the *possibility/necessity* operators of modal logic. Think, for instance, at the fundamental distinction, not only in philosophy, but also in all social sciences known as the "Hume problem", between the *logical necessity* and the *ethical oughtness,* on which the common-sense distinction "facts-values" formally depends. The "social choice theory" consists therefore in the formalization of social philosophy, in strict relationship with the economic philosophy. In a recent paper with the significant title "Logic and social choice theory", Ulle Endriss so synthesizes the core of the social choice theory, as a chapter of formal philosophy:

> When a group needs to make a decision, we are faced with the problem of *aggregating* the views of the individual members of that group into a single collective view that adequately reflects the "will of the people". How are we supposed to do this? This is a fundamental question of deep philosophical, economic, and political significance that, around the middle of 20th century, has given rise to the field of social choice theory [2, p. 333].

The different theories depend then on the *different sets of data* on which the choices are performed and, overall, on the *different aggregation criteria* of individual choices into one only social choice (see below §3.2). Even though the precursor of the theory of social choice is universally identified with Marie Jean Antoine Nicolas de Caritat, the Marquis de Condorcet (1743-1794), and with his famous "paradox of voting" (see below and especially note 4), the social choice theory started in the last century with the pioneering contributions of Abram Bergson and of the Nobel Laureate in Economics, Paul Samuelson leading to the so-called "Bergson-Samuelson SWF" [3, 4], but it received a substantial improvement by the contribution of another Nobel Laureate in Economics, Kenneth Arrow, who, for the first time, gave theory an *axiomatic* form, demonstrating some fundamental theorems [5], and then finally with a further Nobel Laureate, Amartya Sen, who gave the formalism of social choice theory an essential contribution [6, 7] (see below §3.2).

Arrow's SWF is intended as a function ranking social states as less, more, or indifferently desirable, for every pair of them, with respect to individual welfare preferences. In the standard formalism inaugurated by Arrow, we have to do with *linear orders* because based on the standard notion of *transitivity* that is a *linear* relation[1]. Formally, following Endriss [2, p. 338]:

---

[1] Indeed, using the standard *ideographic* symbolism, e.g., the alphabetic letters, *x,y,z,* for representing relations *R*, the ordering relation among objects using the *transitive relation* is

Definition 1. *(Social Welfare Function). Let* $\mathcal{L}(\mathcal{X})$ *denote the set of all linear orders on* $\mathcal{X}$*. A profile* $\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$ *is a vector of linear orders (i.e., preferences), where* R_i *is the linear order supplied by individual i. We write* $\{N^R_{x \succ y}\}$ *to denote the set of individuals that rank alternative x above alternative y under profile* ***R****. (…) A SWF is a function* $F: \mathcal{L}(\mathcal{X})^{\mathcal{N}} \rightarrow \mathcal{L}(\mathcal{X})$ *that maps a profile of preference orders to a single (collective) preference order.*

This strict relationship so instituted by Arrow between the social choice theory and the *order theory* in set-theoretic logic – as Endriss and Sen noticed – historically depends on the fact that Arrow followed in 1940, as an undergraduate student, a course with Alfred Tarski. This justifies the usage of the "lexicographic" (see note 1) symbolism in social choice theory, like in order theory of logic and mathematics. This formalism is characterized by the usage of "alphabetical ordering" or of "numerical ordering" in the symbolism, where, however, numbers are properly "ordinal" and not "cardinal" numbers. From that, the usage of the symbol $\succ$ and derivates ("higher than") instead of $>$ ("greater than") and derivates, since we have to deal here *per se* with "topological" and not "geometrical" algebraic relations, because in principle independent of a given, shared "metrics", on which only a quantitative ordering relation ">" can be defined.

Given this definition of SWF as the starting point of the social choice theory, and emphasized its relationship with formal logic and then with formal philosophy, we can outline the scheme of our contribution.

1.2 The scheme of this contribution

In the **Second Section** of our contribution we illustrate the passage from the SWFs to the *social choice functions* (SCFs), in the framework of social choice theory, in relationship with Sen SCF theory, because the passage from SWFs to SCFs depends largely on Sen's original contribution.

Indeed, another theoretical background to take into account in our examination of the early social choice theories as SWF theories, is their development in the framework of an approach to the study of social and economic systems at *equilibrium* (or in *close-to-equilibrium*) conditions, as the standard formalism of statistical mechanics in physics requires. The so-called "classical" and "neo-classical" economic theories, in which early SWF theories were developed, are all inspired, indeed, by Samuelson's general approach to mathematical economics in his seminal handbook [4], using as reference

---

represented as follows: $((xRy \wedge yRz) \rightarrow xRz)$. As we see, in such a way the transitive relation is *linear,* and then in set-theoretic ordering it supposes a *total ordering* among sets – that is for all sets the ordering relation $\leq$ holds – because admitting *gaps* (in our case, between *x* and *z*) among sets that can be indefinitely long. Another possible transitive relation is the so-called *Euclidean* relation: $((xRy \wedge xRz) \rightarrow yRz)$ that has a *tree structure* in which no gap among sets is admitted, and in which therefore set total ordering is not necessary, but also *partial orderings* are admitted. As we see, this distinction between *total* and *partial* orderings in social choice *rankings* is essential for understanding the novelty of Sen's approach to social choice theory as to the classical ones based on a *liberal* approach in economy.

physical paradigm for modeling economic theories Gibbs' *statistical mechanics,* to which the first two chapters of Samuelson's handbook are significantly dedicated. This is because this mathematical theory is naturally consistent with the liberal individualistic vision of economy and society. In both approaches indeed, statistical mechanics and statistical economy, the *relations among individuals,* either "physical particles" or "economic actors", are considered as *ultimately irrelevant*. And significantly, Sen's criticism to this approach starts just from this point.

Unfortunately, a fundamental unexpected and undesired consequence of Arrow's mathematical theory is the famous "Arrow's impossibility theorem" [5] demonstrating *formally* – to follow the efficacious synthesis of Amartya Sen – the "impossibility of the Paretian liberal" [8] for a consistent social choice theory. In fact, one of the more troubling corollaries of Arrow's impossibility theorem is the inconsistency for democratic systems of social choices based on the majority decisions (see below §2.3) on a liberal basis. That is, any liberal approach to SWF theory like Arrow's one takes as an axiom the so-called "Pareto's efficiency" principle, for which all the individual preference rankings have the very same possibility of determining the social one: an abstract principle that is evidently unrealistic in photographing the effective situation in social sciences. On the other hand, seen from the standpoint of social choice theory, the voting democratic principle is nothing but a way for aggregating individual preferences into one only social ranking using as aggregation criterion the majority principle. The fierce debate triggered by Arrow's results in the next fifty years contributed definitely to the development of the formal and mathematical theory of social choices, as a fundamental mathematical discipline for social, political, and economical sciences. Not casually an impressive testimony of this debate is given by the bibliography of the second enlarged edition, published in 2017 [7], of Amartya Sen's fundamental book *Collective Choice and Social Welfare*, firstly published in 1970 [6].

Now, as we discuss briefly below, one of the most significant contributions of Sen's theory, determining its immediate diffusion among scholars and the wider public, was precisely the formal demonstration that the only way for avoiding Arrow's impossibility results is to enrich Arrow's theory with two new features.

1. Before all, by enriching SWF's ranking (ordinal) measures with (cardinal) measurements and comparisons of resources and utilities among individuals, so to transform SWFs in as many and more realistic social welfare *functionals.*
2. Secondly (and definitely) by introducing into the model the *interpersonal comparison of goods and utilities* – and, more generally, *the information exchanges* among persons constituting homogeneous groups, on the contrary considered as irrelevant in the classical economic theory, as well as in standard SWF theories, like Arrow's one.

In the **Third Section** of this contribution we discuss therefore the main characters of Sen's *social choice function* (SCF) theory. Indeed, the two just remembered improvements as to Arrow's SWF theory allowed Sen to introduce into the mathematical modeling of his SCF theory *comparative principles* of social and economic *equity* or *fairness* using as aggregation principle the ethical *maximin principle* that gives priority to the interests of worst-off persons, used precedingly by his teacher John Rawls in his celebrated political theory of "justice as fairness" [9]. The main difference, as we see, is that, while for Rawls' neo-Kantian *normativism* the maximin principle supposes again – if applied to social choice theory – an untenable in our *plural society* because *absolute* and then *abstract* "total ordering" of the distributive justice grading, in Sen's SCF theory, where only partial orders are allowed, but where it is possible to merge different *local* gradings of justice into one shared ranking. In such a way, the maximin becomes a powerful aggregation criterion of *comparative distributive justice,* not supposing any reference to an absolute criterion, for unveiling and solving concrete injustice situation.

This possibility transforms Sen's SCF Theory into a *comparative theory of social choices*, which considers as *intrinsic* to the models also the ethical and the interpersonal constraints, by Sen's wider and systematic usage of the maximin principle as a *local* equity principle we illustrate below.

On this regard, from the standpoint of the formal philosophy of *ethical choices,* the fundamental contribution of Sen was the formal demonstration that an effective mathematical modeling of ethical constraints in economy cannot be based on abstract and not-computable *optimal choices* defined on a *unique,* because supposed *complete* (=*total*) grading of distributive justice for social/economical states in a society, but on *concrete* criteria of *maximal choices* relative to the different contexts, that is, based on *interpersonal equity criteria*. This mathematical modeling of SCFs can constitutes also the basis for more suitable measures of social well-being and welfare than GDP, recently developed in social, political and economic sciences with the passionate and exciting public discussions that followed, to which Sen gave a fundamental contribution [10].

Finally, in the **Fourth Section** of this paper we want to discuss the computational challenges of Sen SCF theory. Indeed, as Sen himself emphasizes in many places, precisely the difficulty of applying his theory using the classical statistical tools of the economic analysis is the main "technical" obstacle – apart from the ideological ones – against its systematic application. The problem is well synthesized by Sen saying that an approach like his SCF theory all based on the interactions among personal and collective social actors requires from the mathematical standpoint a *dynamic weighing of the variables* that no statistical weighing function can grant in principle. Moreover, from the computer science standpoint, we are faced in machine learning with the very same problem. As we see, indeed, it is absolutely unfitting with a dynamic weighing of the variables the *statistical* machine learning model in AI systems, based on the so-called *deep learning* approach of artificial neural networks (ANN) (see §4.1), maintaining the classical distinction of the statistical analysis, between the determination by the human observer of the statistical weights of the variables, and their further application on the whole dataset. In the ANN modeling this distinction becomes between 1) the necessarily automatized – because of the presence of inner higher order correlations among data that makes impossible the determination of the weights "by hand" – (deep) *learning phase* for fixing the weights of the variables on a representative sample of the dataset, followed by 2) the *testing phase* consisting in the application of this statistical weighing to the whole dataset.

To stress the relevance of this point, let us consider the same problem from another but complementary standpoint. For instance, the real-time information exchange among communicating agents that internet today allows, determines the fast aggregation/dissolution of interest groups in a world-wide environment. For an effective representation of this phenomenon evident to everybody, think, for instance, at our worldwide currency and stock-exchange markets, and at the streams of data, with always changing inner correlations,which internet produces *in real time*[2] before on the computer screens, and then on the consequent choices of the financial operators, distributed over the five continents. As we know, this problem of *infinite* "data streams" in Theoretical Computer Science (TCS) belongs to the wider class of the so-called "big-data" issue, determining the growing interest to innovative computational solutions, the controversial quantum computer architectures included.

Indeed, "infinity" is here related with the continuous, sudden "aggregation/disaggregation" of coherent patterns of behavior of the financial agents (humans and computers), in markets distributed all over the world, "dramatically" represented by the (apparently) erratic oscillations of multiple economical indices on the many screens of each financial operator. The higher order inner, "long-range" correlations among these inter-related oscillating statistical quantities correspond, indeed, to as many "phase transitions" among indefinitely many "phase coherent behaviors" of markets, because of the real time "information exchanges" among operators, both humans and machines.

Coming back to our ANN "deep learning" approach in AI, making "dynamic" the weighing of variables for "locking" the network dynamics onto the continuous data oscillations like Sen's theory, but also like the modeling of data streaming require, would signify to design ANN architectures in *continuous learning* – just as, in parentheses, the natural NNs of our brains are . This, however, makes immediately *chaotic* – in the technical sense of the dynamic system theory notion of "deterministic chaos" – and then absolutely *unpredictable*, using the classical statistical methods, the overall network dynamics.

To sum up, from the precedent discussion it is evident that as fa as we consider the interactions among social and economical actors we cannot any longer consider gas thermodynamics as the reference physical paradigm, but *fluid thermodynamics* having in the quantum field theory (QFT) of *dissipative systems* – that is, "open" systems in continuous interaction with its environment – its fundamental theory. It is not casual that by the principle of the so-called *doubling of the degrees of freedom* characterizing the QFT of dissipative system we can satisfy the requirement of a dynamic determination of the statistical expectations – or dynamic weighing of the variables involved. That is, what we are searching for solving the computational challenges of Sen's social choice theory and/or of data streaming, as far as both characterized by the continuous change of the inner correlations among data.

Therefore, we propose in this contribution the path toward an original solution for this problem, based on the so-called "algebra doubling" of coalgebras by which modeling interactive systems, in computational, and then in physical, cognitive and now also social systems [11]. What is relevant, in our proposal – even though its full

[2] E.g., think at Bloomberg data streaming, updating continuously the values of shares and of currencies on the worldwide financial market, active all over the day because of the different time zones, and that are changing in average every ten seconds.

development would require a wider discussion we cannot develop here – is that such a coalgebraic modal semantics is able to implement in the same computational model, not only the *mathematical component* of Sen's SCF theory, but also its *intrinsic deontic* component based on Sen's theory of "comparative justice as equity" [12]. A theory that is consistent on his revolutionary idea of "development as freedom" [13]. Namely, based on Sen's theory of the fair and concrete distribution of "capabilities" (effectively available choices among different possibilities of "being" and "doing") and of "functioning's" (concrete ways of "being" and "behaving") of persons and groups, in order to fulfil their *free self-achievement as persons*, according to the different value systems they choose to share, in a pluralistic ethical panorama such as ours. On the other hand, the logical consistency and the computational effectiveness of a theory of comparative justice as equity is fundamental for supporting, not only more complex and effective measures of development and of well-being for nations and groups than the obsolete GDP [10], but also for overcoming, by the *formal and then universal comparability* of different moral systems, without supposing an *absolute ethical system of reference,* the moral relativism that negates the *universality of human rights*, and so offends the weaker persons and peoples in our society.

Indeed, by substituting the *abstract* universalism/normativism of a theory of justice, such as, for instance the Kantian one, with a *concrete* and *computationally effective* theory of "comparative justice as equity", because based on criteria of equity among persons and groups sharing different value and moral systems, it is possible to make effectively and universally comparable the *local* solutions of *global* problems for different contexts in our open society. A tremendous task [14] for which the aid of suitable AI systems for supporting the social choices of policy and decision makers in economics and society is unavoidable, according to a notion of "collective wisdom" including necessarily humans and machines [15, 16]. All this emphasizes the relevance of proposing computationally effective models in the field of social choice theory. And these reflections will constitute the *Conclusions* of our contribution.

2 Social welfare functions as social choice functions

2.1 Two main types of social welfare functions

Generally, in economic literature, there are two main types of social welfare functions (SWFs), as far as they are defined respectively, either on

1. Some support (domain-codomain) of real valued economical magnitudes (defined on cardinal numbers) for one only social group, or
2. On orders, i.e., domains-codomains of rankings (ordered sets) of preferences, and of rankings (ordered sets) of social states, rankings defined obviously on ordinal numbers.

Now, when we are speaking about "orders" effectively we are considering a SWF like a particular type of collective choice function (CCF), relating ordered sets of individual preference/utilities, with ordered sets of social/economical states for each individual. This means that, mathematically, we are moving from real valued (cardinal) functions, to set theoretic logic functions – effectively we are moving to set theoretic semantics (order theory) applied to social entities. The initial representatives of these two types of SWF theories generally quoted in literature are, respectively, "Bergson-Samuelson SWF" and "Arrow SWF".

2.2 Bergson-Samuelson social welfare functions

Abram Bergson first introduced in economics the SWF notion as real-valued differentiable functions, aimed at formally representing "the conditions of maximum economic welfare" for the society as a whole. Bergson's SWF then includes as function arguments several quantities of different commodities produced and consumed, and of resources, labor included [3].

The fundamental contribution of the 1970 Nobel Prize in Economics Paul Samuelson – founder of the prestigious "MIT School of Economics" counting among its members an impressive list of Nobel Prizes – is synthesized in what is commonly called the "Bergson-Samuelson SWF". It aims at representing (in the maximization calculus) all real-valued economic measures of any belief system – "that of a benevolent despot, or a complete egotist, or 'all men of good will', a misanthrope, the state, race, or group mind, God, etc." – required to rank consistently different feasible social configurations as "better than", "worse than", or "indifferent to" each other in an ethical sense [4, p. 221].

What is essential for our aims is, anyway, that Samuelson's modelling of the equilibrium stability for economic systems explicitly depends on Willard Gibb's statistical mechanics interpretation of thermodynamic systems, as Samuelson explained in the first two chapters of his masterpiece *Foundations of economic analysis*, of which eighth chapter is dedicated to welfare economics and then to his SWF interpretation in such a theoretical framework [4, pp. 203-256].

2.3 Arrow social welfare function and its "impossibility theorem"

The second type of SWF is related with the work of Kenneth J. Arrow, awarded with the Nobel Prize in economics on 1972. Since the first version of his theorem (1948), indeed, he transformed effectively Bergson's SWF into a *collective choice function* (CCF), as the same title of his fundamental book *Social choice and individual value* exemplifies [5]. That is, whereas the Bergson-Samuelson SWF rules the mapping from any set of individual orderings of preferences/utilities into *one only set of ordering* of social states, Arrow SWF rules the mapping from any set of individual orderings of preferences/utilities into a set of social states, *among many alternative ones* [5, pp. 22-25]. As Arrow himself emphasizes the interpretation of SWF as a CCF is effectively a restriction over SWF, because it requires that for any individual ordering, for some sufficiently wide but finite range of them, the SWF "give rise to a true social ordering" among finitely many ones.

Arrow's CCF effectively poses other conditions to SWF giving rise to the famous "Arrow's impossibility theorem". Following the later version of Arrow's theorem [5, pp. 46-61] these conditions can be synthesized as follows.

Let A be a set of alternatives of social states, N a number of individuals or preferences i, and $L(A)$ the set of all linear orderings of A. The SWF, intended as an individual preference aggregation rule, is a function F: $L(A)^N \rightarrow L(A)$ which aggregates *individual preferences* R_i *on* N into a *single order* on A. The N-tuple $(R_1, \ldots, R_N)$ of R_i is called an "(individual) preference ordering". The theorem states that for N at least of 2 individuals, and for A at least of three alternatives the follow four conditions *are incompatible* – particularly the fourth one as to the others, and particularly as to the second one, as Sen will demonstrate:

1. U^*: *Unrestricted domain*. The domain of the function (rule) F must include all logically possible individual orderings for a finite set of them;
2. P^*: *Pareto efficiency or "unanimity"*. All individual orderings have the same possibility of determining the social state ordering, i.e., if alternative x is ranked strictly higher than y for all orderings $R_1, \ldots, R_N$, then also for $F(R_1, \ldots, R_N)$;
3. I^*: *Independence of irrelevant alternatives*. F depends strictly on the pairwise relations associating subsets of A and N, i.e., if for two different individual preference orderings R and S the alternatives x and y have the same order in R and R', then also in $F(R_1,\ldots, R_N)$ and $F(S_1,\ldots, S_N)$.
4. D^*: *Non-dictatorship*. There is no individual $i \in \{1,\ldots N\}$ such that for $\forall(R_1,\ldots, R_N) \in L(A)^N$, x ranked strictly higher than y implies that x is ranked strictly higher than y, for all x, y.

To sum up, the "Arrow impossibility theorem" states that "there is no SWF satisfying simultaneously conditions U^*, P^*, I^*, and D^*". Roughly speaking, given the condition P^*, for any individual set ordering, there must be one *pivotal* individual that orders the alternatives. Now because of the transitivity and asymmetry of the *strict* individual orderings supposed in Arrow's construction (see below), it is possible to demonstrate that this pivotal individual is *unique* for all individual orderings, so violating the condition D^*.

The theorem had during the last fifty years a tremendous impact overall in political sciences because, if taken for granted, it would demonstrate the irrationality of vote systems in democracy, as far as based on the principles of "majority" and of "representativeness", which suppose evidently that the conditions P^* and D^* *hold simultaneously*[3].

Effectively, two are the main ways for avoiding the outcome of Arrow's theorem:

[3] There exists evidently, and many Authors noted it, a relationship between Arrow's impossibility theorem and the famous *Condorcet's paradox on voting* stated in his *Essai sur l'application de l'analyse à la probabilité des decisions rendues à la pluralité des voix* published in 1785 [48], and considered unanimously as the official birth-date of the modern, mathematical theory of social choices. However, Sen demonstrated that Arrow's impossibility theorem cannot be reduced to Condorcet's paradox, before all because "it would be hard to argue that majority rule would really be 'a plausible way of aggregating preferences' in welfare economics" [7, p. 275]. See also [7, pp. 395-419], and overall the concluding chapter of Sen's most recent book [7, pp. 453-472], in which he offers a deep analysis of what a modern comprehensive, because rational theory of "social choice" requires, and that is far beyond any oversimplified and oversimplifying "majority criterion".

1. Working on *infinite sets* of preferences – i.e., on finite sets of *ever-changing* preferences and aggregation rules –, and not on finite (*static*) ones like in Arrow's SWFs. In such a case, however, as far as we suppose a *total ordering* relation of sets, the aggregation rules seem to be of limited interest because in standard set-theoretic logic they are based on *ultrafilters* – i.e., *maximal* filters on given "partially ordered sets" or *posets* (see Rel. 9 below), supposing the *total ordering* (see Rel. 8 below) of all partial orders. Namely, this means a supposition of *completeness* in the definition of the social rankings that is simply illusory to suppose in complex and global societies like ours. Ultrafilters, in other terms, are *non-constructive* mathematical objects because they suppose (infinite) set *total ordering* – i.e., their definition requires a higher order logic –, so to suppose practically as many "invisible dictators" [17]. And, in fact, such aggregation rules violate Turing computability [18], so to result *practically useless* in mathematical social choice theories and in social computing applied to economics and to social sciences.
2. Working on set *partial orders,* instead of the *strict-orders* of Arrow's SWF (see below), so to define another type of rational CCFs, the *social decision function* (SDF). For SDFs Arrow's impossibility theorem, indeed, does not generally hold, even though it holds for SWFs as special cases of SDFs. This approach was firstly proposed by another Nobel Prize in Economics (1998), Amartya Kumar Sen, who gave in this way an essential contribution of clarification to our discussion [6, 7].

We can anticipate, however, that also in Sen's approach the "challenge of infinity" and hence of effective computability on infinite sets, e.g., data streams, represents itself as an unsolvable computational problem within the Turing Universality paradigm in computations. This holds also in the case of the "probabilistic" Turing Machine architecture, *neural networks* included, as far as based on a purely *statistical weighing* of the variables, through the machine learning principle – the so-called "deep learning" of actual AI systems – of the *stochastic gradient descent,* as we discuss briefly in the **Third Section** of this contribution . At least, till we are working in standard set theory – e.g., Zermelo-Fraenkel (ZF) set theory – and not in non-standard ones, like in the theory of Non-Wellfounded (NWF) sets requiring Algebraic Universality in computations, and not the Turing Universality, as we hint below.

2.4 Amartya Sen's logic of preferences and his notion of "social choice function"

Formally, Sen's contribution to the discussion consists in a rigorous clarification of what "preferences/evaluations among alternatives" (e.g., "better than", "as good as", "at least good as", etc.) means in the framework of the *logic (algebra) of relations*, for solving ambiguities and inconsistencies that often appear in SWF theories interpreted as CCFs. We illustrate Sen's contribution essentially following his outstanding book *Collective choice and social welfare,* in its first edition published in 1970 and then in 1979 [6], in which he outlines a mathematical logic of preference/utility in economic and social sciences. As we anticipated in the Introduction, Sen recently published in

2017 an enlarged edition of this fundamental work including the first one, plus some further extension we discuss below [7].

To summarize briefly the core of a formal approach to social choice theory, we have to recall the basic fact that mathematically whichever preference/evaluation binary relation R over a set S consists in specifying a subset R of the Cartesian Product $S \times S$, defined as the set of all ordered pairs (x, y) such that $\langle (x, y) \in S \rangle$.

Following the standard order theory in set theoretic logic, Sen defines rigorously the following relations R over a set S in order to offer a *lexicographic* theory of social orderings:

1. *Reflexivity*: $\forall x \in S: xRx$
2. *Totality* (*Completeness*): $\forall x,y \in S: (x \neq y) \rightarrow (xRy \vee yRx)$
3. *Transitivity*: $\forall x,y,z \in S: (xRy \wedge yRz) \rightarrow xRz$
4. *Anti-symmetry*: $\forall x,y \in S: ((xRy = yRx) \rightarrow x = y) \vee (xRy \neq yRx) \rightarrow x \neq y)$
5. *Asymmetry*: $\forall x,y \in S: xRy \rightarrow \neg (yRx)$
6. *Symmetry*: $\forall x,y \in S: xRy \rightarrow yRx$

For logical and computational aims, it is worth to notice that anti-symmetry includes asymmetry, and not vice versa, so that the implication connective is logically correct only in the case of anti-symmetry. Indeed, in the case of equivalence – which is a double, symmetrical implication – anti-symmetry but not asymmetry holds, since asymmetry holds in the case of simple implication only. This observation is essential for understanding the logical notions of *partial* and *total orders,* in all their possible versions, where anti-symmetry plays an essential role. Effectively, Sen defines furtherly, and newly in a standard way, different types of "ordering relations", according to the different 1.-5. basic relations they satisfy:

7. *Pre-ordering*: 1, and 3;
8. *Total* (*Complete*) *pre-ordering*: 1, 2, and 3;
9. *Partial ordering*: 1, 3, and 4;
10. *Total (Complete) ordering*: 2, 3, and 4 (in the supposition that 2 implies 1);
11. *Strict ordering*: 2, 3, and 5.

Sen defines thus the ordering relations 7, 8, and 9 as "quasi-orderings" that is the key-notion of his logic of preferences, evidently based on the notion of "pre-orders" [6, p. 10]. Then, Sen notices, in the case of Arrow's SWF, the "ordering" must effectively satisfy 1, 2, and 3, but it is, *per se,* irrespective of 4 and then of 5 [6, p. 8]. In such a way, we can avoid, under some conditions explicated by Sen in the rest of the book that we only summarize here, the *uniqueness* of the "pivotal individual" for all social orderings, as depending on some rational CCF on social states he defined as "social decision function" (SDF). This opens the possibility of satisfying the conditions P^* and D^* simultaneously of Arrow's theorem as characterizing any CCF, and then a SDF too, even though at the price of an organization of social states much more "fluid", than in Arrow's vision.

In other terms, the core of Sen's approach consists in demonstrating formally that we can obtain suitable CCFs such as SDFs, satisfying all Arrow's conditions without falling into his “impossibility result”, by supposing *quasi-orderings* and overall *partial orderings* instead of *strict orderings.* We refer to [7, pp. 313-337]**,** for a complete discussion of Sen’s formal treatment of the CCFs and SDFs in social choice theory.

To conclude this subsection, let us recall, another interesting consequence deriving from the *finiteness* of SDFs – and of the social choice functions (SCFs) we discuss below.

That is, the possibility of representing in it also *judgements about values,* shared by a homogeneous social group because characterized by strong *interpersonal exchanges of information.* These topics were the main object of a more recent essay of Sen, with the significant title: “The Informational Basis of Social Choice” [19], and they are the *leit motiv* of the new part added to the Second Expanded Edition of [6], even though largely discussed also in [6, pp. 89-117], where he introduced the transition from a SDF to a *social choice function* (SCF). We return in the next Subsection on this fundamental transition.

However, when we study social and economic systems, taking into account also the *motivational social forces* acting within them, overall if vehiculated in *real time* by social media, the *challenge of infinity* represents itself because making untenable that either a SDF or a SWF can range over a *fixed finite set* of social state alternatives, chosen by *individual* preference orderings.

Nevertheless, the situation changes radically when we pass from individual preference orderings (rankings), to *social* choice functionals, because expressed by *homogeneous groups of persons*, that is, by groups of individuals characterized by a *strong mutual information exchange* making them an unique *social subject*: this is the core of the passage from SDFs, to SCFs.

Using a thermodynamic metaphor, this is the core of the passage from a liberalist vision of society seen as a “gas” or “vapor” of individuals, to a “pluralist” vison of society characterized by the “condensation” of individuals into different homogeneous groups – like as many “drops”, for continuing in the thermodynamic metaphor. In such a case, the “infinity challenge” represents itself in a different way, because of the “boundary condition” of the social media influence, often on a world-wide dimension – think at internet, before all. The influence of social media indeed, in specific contexts, acts like sudden changes of “temperature” in the human environment, determining immediate rarefaction/condensation processes of old/new groups inside the human community, often in a world-wide, super-national dimension. In a word, such a “liquid” situation of our society and of our economy requires a “paradigm shift” in the mathematical approach to economic, social, and political sciences. That is, a passage to a modelling based on *fluid thermodynamics* that at the fundamental physical level is modeled in the quantum field theory (QFT) of dissipative systems as we discuss briefly in the **Fourth Section,** being the common coalgebraic modeling, the formal bridging between mathematical physics and mathematical logic.

3 Sen's notion of social choice function: its background

In this subsection we develop our analysis in several steps. After a general definition of what is a social choice function (SCF) in social choice theory, before all, we examine Sen's social choice function definition, and the consequent "SCF possibility theorem".

Afterward, we discuss the relevance of the information basis in SCFs.

Furtherly, we examine John Rawls theory of "justice as fairness" in the context of social choice theory because this prepares Sen's theory of "justice as equity", on which his SCF theory is based, and that constitutes its own originality and complexity.

The examination of this theory constitutes the object of the next subsection §3.4. Indeed, Sen's SCFs theory strictly depends on his original notions of the personal "capabilities" and "functioning's", as the key-concepts of his theory of "justice as equity".

3.1 "Sen's SCF and "SCF Possibility Theorem"

Generally, as we have seen, SCF in axiomatic social choice theory is a function $F: \mathcal{L}(\mathcal{X})^{\mathcal{N}} \rightarrow 2^{\mathcal{X}} \backslash \{\emptyset\}$ mapping profiles of linear orders $\mathcal{L}$ on alternatives $\mathcal{X}$, to non-empty sets of alternatives.

> Intuitively, given a profile of declared preferences, *F* will choose the 'best' alternatives" [2, p. 341].

Therefore, according to Sen, the possibility of eschewing the notion of individual or social preferences, toward a theory of *social choice functions* "received considerable attention in the literature of social choice theory in recent years" [7, p. 313]. Indeed, in the case of SCF theory, it is systematically avoided the necessity of defining a set of binary preference relations over *one only* ranking (ordering) of alternatives. A non-binary formulation of social choice, indeed, works directly with a choice function *C* for a society, defined over the alternative rankings that can be chosen by individuals and by groups, i.e., over a choice set *C*(*S*) for each menu (set) of alternatives *S* (for the formal definitions of a collective choice function (CoCF) over a *C*(*S*).

From this Sen's *general* definition of a SCF follows [7, p. 313]:

Definition 2. *(Definition of Social Choice Function). The functional relation between the set of alternative menus S and the corresponding "choice set" C(S) can be called the social choice function (SCF), and this needs not invoke (…) any binary relation of social preference.*

A "functional collective choice rule" (FCCR) maps *n*-tuples of individual preference relations R_i to a social choice function:

$$C(\bullet) = f(\{R_i\}) \tag{1}$$

It is evident that a SCF escapes from Arrow's impossibility results simply because Arrow conditions are essentially *binary* relational demands (i.e., individuals *vs.* one social state ranking). In fact, in the present case, what is required is not a SWF (or SDF) (see §2.4), but a SCF:

$$g\left(S,\{R_i\}\right) \text{ such that } g\left(S,\{R_i\}\right) \subseteq S \text{ for all non-empty subsets } S \subseteq X \tag{2}$$

This is essentially equivalent to making the value of the function $f(\{R_i\})$ a *finitely complete* choice function $C(\cdot)$ for each society, and not – as with SWFs or SDFs – a social preference relation *R* [7, p. 315].

In this way, for such a *functional collective choice rule* or FCCR: $f(\cdot)$, the SCF theory requires a re-definition of Arrow's conditions *U, P, D* and relative axioms (see §2.3), in terms of restrictions of *n*-ary choices over pairs only – namely, $\widehat{U}, \widehat{P}, \widehat{D}$ plus the extension of non-dictatorship condition *D* to a stronger condition of *anonymity* $\hat{A}$ ("if $\{R_i\}$ is a permutation of $\{R_i'\}$, then $f(\{R_i\}) = f(\{R_i'\})$" – that we cannot discuss here, and for which we refer to [7, p. 316]. Now, if we combine these conditions to the other Arrow's condition *I* of "independence of irrelevant alternatives" that was defined in functional terms, i.e., as an I^A, by Arrow himself, we can state the following theorem [7, p. 316]:

Theorem 1. *(Choice-functional positive possibility theorem). For #H≥2, there is a FCCR satisfying conditions* $\widehat{U},\ I^A, \widehat{P}, \widehat{D}, and\ \hat{A}$.

Where, in the following discussion [7, pp. 316-336], Sen emphasizes the relevance of having an *ariety* (number of arguments) *#H* ≥ 2, where *H* is the set of individuals, in a social choice functional interpretation of SWFs and SDFs, since, also in the case of restrictions of choices over pairs only, a FCCR must satisfy anyway the condition of "transitive closure" xB^*y of a sequence of at least a triple of binary relations *xBy,* having necessarily a "maximal" choice set including the factor B^A of *B,* defined as *xBy* and *not yBx,* that is an asymmetry condition, which closes the sequence, avoiding any cyclicity. We return in §3.4 on the centrality of this *maximality* condition for emphasizing the computability and then the *effectiveness* of Sen's SCF theory also with respect to its *ethical personalistic components.*

Anyway, this can be seen straightforwardly also in the more elementary case of *Condorcet's paradox of voting*, interpreted as a choice function. In fact, the *transitive closure* intrinsically associated with a maximality condition avoids the cyclicity of preference orderings, from which the paradox derives[4]. In a word, it is the *finite completeness* of a choice function, because including necessarily an *intrinsic maximality* criterion,

[4] Indeed, the paradox derives from the fact that voter 1, 2, and 3 are ranking the three alternatives in the orders: *x,y,z*; *y,z,x*; and *z,x,y*, respectively. In this way, there is a preference cycle, and each alternative is defeated in a majority vote by another alternative. Now, if we do not consider *preference maximization,* but *maximality* in terms of the *transitive closures* of preference relations, we have the following choices: $C(\{x,y\}) = \{x\}$; $C(\{y,z\}) = \{y\}$; $C(\{z,x\}) = \{z\}; and\ C(\{x,y,z\}) = \{x,y,z\}$. Of course, this choice function is not binary, but satisfies all the Arrow axioms in their choice-functional version, of which at **Theorem 1** above (see [7, p. 316]).

the core of a SCF theory. And this, without supposing, either the maximization of some *extrinsic* "optimization function", or some criterion of *internal* "consistency" among the different menus of social alternatives that are, on the contrary, independent of each other, as it is the case of a "plural society" such as ours.

3.2 SCF theories and their information bases

However, it is not in this possibility theorem extended also to a theory of voting and then fundamental for a formal justification of a democratic regime based on the majority principle, the main advantage of Sen's SCF with respect to Arrow's SWF (see on this point [7, pp. 220-250; 395-419], and below §3.4).

Indeed, more generally, what is systematically lacking in Arrow's social theory is the possibility of an *interpersonal comparison* of people well-beings, using several different quantitative indices, or of any other *comparative data* about people's respective advantages. Effectively, as Sen noticed, Arrow was following in this the general consensus emerged in 1940s among scholars that "the interpersonal comparison of utilities has no meaning" [5, p. 9], and that is formally related to a "gas-like" or individualistic (liberalist) social theory, as we said many times. On the contrary, if it possible to allow in a mathematical social theory the comparison of social states based on "interpersonally comparable information about different persons, many possibilities of making systematic social welfare judgements open up". Particularly, we can insert in the social welfare theory "*ethical concerns* such as equity, or removal of deprivation, or enhancement of people's freedoms, etc." [7, p. 338].

For this aim, in the light of the above equation (1) defining a SCF, Sen examines, in a very original and intriguing way, different theories of social justice in terms of their "informational bases". That is, in terms of the presence of information exchanges about states, achievements and opportunities of the people, which are central to assessing justice and injustice and their grading in a given society according to the examined theory, at the same time ruling out information of other types considered as irrelevant.

For instance, the *utilitarian theory* of justice attaches relevance to the economical utilities of individuals, and it does not consider information about fulfilment of violation of rights or liberties, levels of incomes or affluence that people enjoy – if not for the indirect effects that such variables have on individual utilities. On the contrary, a *libertarian theory* of justice stresses over individual liberties and rights, and it gives no direct relevance to the levels of utilities or of incomes, and of their interpersonal comparison.

To sum up, Sen defines the informational bases of different theories of justice for inserting them in the functional structure of a SCF in terms of [7, pp. 339-340]:

1. The *basal space* of the theory, i.e., the class of variables on which the justice assessment of that theory depends (e.g., in utilitarian theories, the combinations of the individual utilities);
2. The *aggregation system* of those variables, for allowing a discriminating use of these variables (e.g., in utilitarian theory, individual utilities are simply added together to arrive at a sum-total serving as basis for the assessment of the social state, without

paying attention, for instance, to some measure of dispersion or inequality). In this context, the aggregation systems of theories of justice including as their essential components "distributive criteria" of commodities, utilities, and opportunities among individuals – that is, the so-called *theories of justice as equity* like Rawls' and Sen's ones –, are particularly relevant.

3.3 SCF theory and Rawls' theory of "Justice as Fairness"

Effectively, in all his work and not only in this book in both its editions (see [6] Chapters 9 and 9*, [7, pp. 187-219; 347-364]), Sen pays a particular attention to the theory of *Justice as Fairness* of one of his former teachers, John Rawls [20, 9, 21, 22, 23].

This theory influenced deeply not only Sen's theory, but the political philosophy of XX cent. overall in US, in the sense of a "moderate liberalism", if not in the sense of a "social-democratic liberalism", even though such labels are not able to do justice of the relevance, the richness and the originality of this theory.

Indeed, Rawls' theory of "justice as fairness", because it introduces a fundamental criterion of "equity" or "fairness" – that is, of *interpersonal comparison of welfares* – as constitutive of the idea of "justice", establishes also a substantial point of distinction as to the Kantian *formalist* theory of justice. Even though Rawls wants to offer by his theory an updated version of Kant's ethics, in the direction of a theory of *distributive justice,* without losing the *unconditional* character of the moral principles characterizing Kant's theory.

Not causally, indeed, in his systematic treatment of Rawl's theory in both editions of his book on *Collective Choice* [7, pp. 187-209], Sen starts precisely from the quotation of the fundamental statement of I. Kant's *Grundlegung zur Metaphysik der Sitten* (1785) [24], synthesizing his formal principle of "duty for duty" as the *unconditional* basis of the "moral oughtness", as distinct from the *unconditional* "logical necessity". Namely, this principle reads: "act always on such a maxim as thou canst at the same time will to be a universal law"[5]. As we know, in this way Kant wanted to satisfy the so-called "Hume's law" distinguishing between *logical necessity* and *ethical oughtness* – sometimes denoted also as the reciprocal irreducibility "facts"/"values" – in philosophical (modal) logic, at the same time preserving the *universality* of moral laws by granting them the same *unconditional* or *absolute* character of the logical laws – that is, by granting to moral laws the same tautological character of logical laws by the principle of the "duty for duty" – without confusing them. Historically, in this way, Kant wanted to criticize the Aristotelian foundation of justice on equity made proper by the Scholastic Tradition, which seems to imply a factual comparison of what is *effectively* right for different individuals in similar circumstances.

Not casually the modern re-proposals of the theory of "equity" or "fairness" before Rawls, as Sen recalls, tried systematically to reconcile the notion of equity with the "absolute universality" of the Kantian moral law, by making irrelevant the personal differences. On this regard, Sen quotes Henry Sigdwick's approach that is paradigmatic

[5] Quoted in [7, p. 187] from [49, p. 66]

in this sense because perfectly compatible according to him with Kant's principle [7, pp. 187-188]

> … Whatever action any of us judges to be right for himself, he implicitly judges to be right for all similar persons in similar circumstances. Or, as we may otherwise put it, 'if a kind of conduct that is right (or wrong) for me is not right (or wrong) for someone else, it must be on the ground of some difference between the two cases, other than the fact that I and he are different persons.' A corresponding position may be stated with equal truth in respect of what ought to be done *to* – and not *by* – different individuals [25, p. 374].

In the formalization of Sigdwick's theory of equity firstly proposed by Richard M. Hare [26, 27], the universality of moral judgements consists ultimately in inserting their "invariance for permutation of persons" within the "similar circumstances" evocated by Sigdwick, everything else remaining the same. Of course, Sen notices, there was a lot of criticisms to Hare's modelling of equity theory as a formal basis of a social choice theory, because it requires that a moral judgement be held from *every* position that a person can occupy in a society through interpersonal permutations. The reference to Sidgwick's and Hare's theory of equity is useful to Sen, however, for introducing the originality of Rawls' theory about justice as equity, for which no interpersonal permutation of positions is required since for Rawls "the principles of justice are those which would be accepted in a fair situation in the *original position*" [7, p. 192].

In a paper of 1958 that anticipates his famous book of 1971, but it is often ignored by the scholars, Rawls defined, indeed, the *original position* as "an initial position of equal liberty (which is defined by the pattern of rights and duties, powers and liabilities, established by a practice)" [20, p. 166], from which any theory on the principles of justice must start. In Sen's words, "the principles of justice are those which would be accepted in a fair situation in the *original position*" [7, p. 192].

Effectively, this position is for Rawls a purely *hypothetical* and *ahistorical* situation – and in this sense it is not comparable with J. J. Rousseau's primordial "state of nature" of humans. It is "hypothetical" in the sense that the principle to be derived from the original position are what the individuals *would agree to*, under given legitimating conditions, not what they *have agreed to*. It is "ahistorical" in the sense that it is not supposed that such an agreement has ever been or could ever have been. In other terms, the "original position" is where individuals choose "principles" in a state of basal equality without any reference to their own place in different social states [9, 23]. In this situation, these principles would satisfy the criterion of "fairness", being the result of a fair agreement with no vested interests, also because of a "voluntary veil of ignorance" of individuals about their *actual* personal differences. In Kripke's modal logic jargon of "possible worlds" semantics applied to deontic logic (see below, **Section 4**), the Rawlsian "original position" is therefore an "ideally good possible world" that never can be our "actual world".

To sum up, the main *two principles of distributive justice* deriving from the "original position" are for Rawls the following [9, p. 266]:

> **First Principle**: Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all. **Second Principle**: Social and economic inequalities are to be arranged so that they are both: (a) to the greatest

> benefit of the least advantaged, consistent with the just savings principle, and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.

These two principles are interpreted by Sen as constituting the “information basis” of Rawls’ theory of justice. That is, respectively, as determining the “basal space” of the main variables characterizing the theory (First Principle), and as determining the “aggregation criteria” of these variables (Second Principle), mainly the *maximin principle* (a).

In this framework, according to Sen [7, p. 348ff.], the “basal space” of Rawls’ theory is constituted essentially by the “basic liberties”, but in reference to the “primary goods” that each individual as citizen must hold, and this is a first point in which Rawls differentiates himself from the libertarianism. Indeed, Sen emphasizes, the range of these liberties is narrower than in libertarianism – e.g., “it does not include property rights or rights of exchange or of bequeathal”. Primary goods are, in fact, “general-purpose *resources* that are useful for the pursuit of objectives that the individual may have” so that, in Rawls’ words, they are “things that citizens need as free and equal persons, and claims to these goods are counted as appropriate claims” [28, p. 257].

“Primary goods” are then “things that every rational man is presumed to want”, and include, without any pretension of exhaustiveness, beside “the basic liberties”, “income and wealth”, “freedom of movement and choice of occupation”, “equality of opportunities”, etc. The reference to human rationality as characterizing all human individuals, makes evident that, in moral philosophy, Rawls’ “primary goods” correspond to moral goals, which are proper to humans as such, and then that each individual universally (naturally) wants, so that each person has the right to pursue them in whichever society. Namely, they correspond to what in political and law philosophy are defined as “human rights” that naturally must have their place in any social choice theory (see [7, pp. 420-452]).

On the other side, the fundamental “aggregation principle” of Rawls’ theory is, according to Sen [7, p. 350], the *maximin principle*. Indeed, Sen emphasizes [7, pp. 193-194], the component (a) of the “Second Principle of Justice” in the “original position” we quoted before and by Rawls himself defined as the “Difference Principle”, since it requires that “social inequalities be arranged to make the worst-off best-off”, is essentially a *maximin criterion.* In fact, it requires that the minimal element in the set of individual welfares is maximized.

Now, despite Rawls main focus is on the type of institutions to be chosen in order to satisfy at best his two principles of justice, and not to propose a social choice theory, nevertheless the maximin principle could be used in principle also to order social states based on individual orderings, and so within a social choice theory framework. For instance, for any social state, we may order the individuals in terms of their welfare and pick on the worst-off individual. This can be compared with the welfare of the worst-off individual in another social state, and so on, so to have *recursively*, at least in principle, a *complete social ordering* [7, p. 193][6].

---

[6] Effectively, Sen offered in [6], since its first printing in 1970, a lexicographic formal version of the maximin principle that Rawls himself included in his *A Theory of Justice* of 1971 [9].

Nevertheless, Sen rightly emphasizes, the maximin principle interpreted as a sort of choice procedure in a social choice theory is not a SWF in the sense of Arrows, since the maximin procedure imposes that a worst-off individual *i* in a social state *x* goes up for every alternative so to change his situation. This means that the social ordering involving *x,* being based on a different individual's welfare, becomes different. And this is incompatible with any SWF [7, p. 193].

Neither the maximin procedure is a collective choice rule (CCR), since a CCR specifies a social preference relation based on the set of individual orderings of social states, each from one's own point of view. In other terms, "a CCR is based on *n* orderings of *m* elements, whereas a Rawlsian maximin choice mechanism is based on one ordering of *mn* elements" [7, p. 194].

At this point many problems raise in the possibility of interpreting Rawls' maximin principle as an effective *decision rule* in a social choice theory – as unfortunately it has been done too many times by policy makers, and in recent times, surely unintentionally because of their generalized and proudly vindicated lack of competence, by the so-called "populist movements". Indeed, if it is absolutized as a decision rule in social choices, the maximin criterion becomes easily a factor of *social injustice* and *economical regression*, given that it oversimplifies complex situations. And in fact, it is often accompanied by misinformation campaigns on the media, which are, in fact, a perverse way of implementing the Rawlsian principle of the "veil of ignorance" over the real causes of the actual differences, making this veil not "voluntary" but "imposed".

Anyway, some of the main limitations of using the Rawlsian maximin principle as a decision rule in a social choice theory are summarized by Sen in [7, pp. 195-198]. Ultimately, they reduce themselves to the fact that Rawls' criterion of "fairness", as far as referring to a *hypothetical* "original position" that, for definition, is irrespective of the historical, *actual* personal differences (covered by a "voluntary veil of ignorance"), becomes a too severe condition – in some proper sense, "ascetical", because this requires that individuals renounce voluntarily to take into account their personal differences –, for being used in a theory of decisions about *effectively fair social choices.* Indeed, we repeat, it risks becoming a false strategy of curing the symptoms of an illness without removing the causes – or because there is the precise will of not removing them, as it is the case of the social media manipulators behind populist movements. Nevertheless, it remains true, as Sen emphasizes, that in several institutional questions that is the proper focus of Rawls theory of justice, the appeal of the Rawlsian maximin criterion is well demonstrated and it works. Precisely because it cures one – if not the principal one – cause of a social disease: the existence in a society of some unjust institutions, because not aimed, each in its proper way (e.g., as an educational, or as an occupational, or as a political institution, etc.) at maximizing the social states of worst-off individuals.

However, what Sen notices is that the maximin principle suggested by Rawls can be used also independently of his theory of justice, in a social choice theory. Particularly in a theory of "justice as equity" like that implemented in Sen's SCF theory (see below §3.4). It, indeed, starts from the actual *personal – also subjective and not only objective – differences among individuals,* the situations of *manifest injustice or iniquity,* without "veiling them", before all. That is, it starts from the relevance in a theory of justice as equity of the *subjective* differences (including the ethical and cultural differences, the religious beliefs, or even the simple personal tastes) in the *interpersonal* comparison of welfare states, because determining concretely different *subjective* ways and grades of enjoying what are, *objectively,* the same commodities, and/or of profiting by what are *objectively* the same utilities.

Indeed, precisely such differences among individuals and groups in a society make untenable the "Pareto unanimity axiom" we discussed in §2.3 so to transform Arrow's "impossibility theorem" in Sen's "impossibility of the Pareto liberal" [8], because incompatible with the definition of equitable social rankings[7]. Indeed, to sum up the example discussed at length by Sen in [7, pp. 207-209]**,** if the increasing of GDP in a given nation, e.g., in India, depended on the growth of production of *beef*, the impact would be different over Muslims and Christians welfare states, with respect to Indus one. Just as, on the contrary, it would happen for the welfare states of Indus and Christians, as to Muslims one, if GDP increasing depended on the growth of production of *pork*.

In a word, in Sen's theory of justice as equity, the interpersonal comparison of states and preferences for defining a common ground of equality, on which applying the maximin criterion, is ethically based not on a voluntary "veil of ignorance" over the differences, but on the ethical "principle of 'extended sympathy' in the form of placing oneself in the position of another" [7, p. 290], starting from an *extended* individual ordering $\tilde{R}_i$ of persons *i, j,* and states *x, y*, (*x, i*), (*y, j*), where $i \neq j$, that is, where they are not mere individuals *i*, but *persons* i.e., "individuals-in-relation".

This principle of "extended sympathy" is formalized in Sen's SCF theory by using two "axioms of identity" among individuals (persons), the first of which reads: "each individual *j* in placing himself in the position of person *i,* takes on the tastes and preferences of *i*" [7, p. 214]. This ethical *subjective* principle can be *objectively* extended to a second axiom of "complete identity", so to give a formal criterion of *grading justice* for the "interpersonal comparison and permutation" among social states of different persons in a SCF theory. That is, the first axiom can be extended to the identification between different *preference rankings J* of the same social states, *x, y*, for any pair of different persons *i, j,* i.e., $xJ_iy = xJ_jy$, where xJ_iy reads: "*x* is more just than *y* according to person *i*" [7, p. 210]. What is fundamental for the logic and the computability of Sen's SCF theory, is that, while each J_i in P. Suppes' theory of grading justice [29], from which Sen derived its own, is a *"strict" partial ordering* over the set *X* of possible social states, if we insert the possibility of making equivalent two or more of them originally different, each J_i becomes a *"simple" partial ordering*[8] O_i, which is anyway

---

[7] Effectively, in the handbooks of social choice theory Sen is often recalled essentially for his criticism of the "Paretian liberalism" (see, for instance, [50, pp. 844-847]), synthesized in his famous theorem of "impossibility of the Paretian liberal", presented in the Chapters 6* of [6], since of its 1970 edition, and discussed at length in the expanded edition of 2017, after the fierce debate it ignited during almost fifty years (see [7, pp. 139-141; 420-452].

[8] We recall here what anticipated in §2.3. Namely, that a "strict" partial ordering is an ordering R $(\leq)$ among sets, in which only the *asymmetric* case $(xRy \neq yRx) \rightarrow x \neq y$ of the *antisymmetric* relation that holds for a partial ordering as such. That is, where also the *symmetric* case $(xRy = yRx) \rightarrow x = y$.

sufficient for extending the Rawlsian maximin principle as an aggregation principle to Sen's SCF theory (see Chapter 9* of his book in both its edition [7, pp. 210-220]).

In a word, what in Rawls justice theory is a *hypothetical condition* of fairness among individuals, hiding under a voluntary "veil of ignorance" the objective and subjective *actual* differences among *persons*, in Sen's theory such differences are the *starting* point, so that fair judgements and solutions of unjust inequalities become the *final* point of a process of interpersonal comparability and information exchange, in the limit leading to an ethos of *dialogue* among different persons and groups, in a multi-cultural society.

Complementarily, for any assumption of real-valued (statistical) measurability and interpersonal comparability of personal welfare states, in a SCF theory based onto Sen's theory of "grading justice" as interpersonal equity, the axiom of complete identity becomes operationally an equivalence between vector spaces *V, V'* of *mn* alternatives. That is, of *m* states for *n* individuals for two different persons *i, j*, onto which calculating the statistical expectations in a quantitative theory of *real-valued* welfare functions (see again [7, pp. 210-220] for the formal treatment of these notions).

To conclude this Section, let us discuss briefly what is he proper of Sen's theory of justice as equity among different persons, characterized by different "capabilities and functioning's", intrinsically related, *objectively,* to different social contexts and related opportunities, and, *subjectively*, to different goals and value systems, as a fundamental *ethical* component of Sen's SCF theory.

3.4 Sen's theory of social choice function based on his comparative theory of distributive justice

Sen's theory of personal "capabilities and functioning's". What characterizes Sen's theory of justice [12, 13] and of his many followers (see, for instance [30, 31, 32, 33]) as to the classical liberal and utilitarian theories, on the one side, and as to Rawls theory of justice as fairness, on the other side, is synthesized by Sen himself in the following passage

> In recent years there has been considerable discussion on an approach to justice that concentrates on people's capability to lead the kind of life they have reason to value – the things that they can do, or they can be. The roots of the approach can be traced to the ideas of Aristotle, and, to some extent, Adam Smith; it concentrates on the opportunities that people have to lead valuable and valued lives. Aristotle saw this achievement in terms of 'human flourishing'. Among other things, he pointed out, in *Nicomachean Ethics*, that wealth 'is evidently not the good we are seeking' – 'for it is merely useful for something else' [7, p. 356].

In other terms, the core of Sen's theory of justice as to liberal and utilitarian theories consists in considering commodities and utilities as *means* for reaching those *goods* or "valuable ways of being and of doing", which a person – on the basis of her/his system of values – considers as fundamental for living "a valuable and valued life", that is, a *personal achieved life.*

> A personal achieved life can be seen as a combination of "functioning's" (i.e., of doings and beings), and, taken together, can be the basis for assessing that person's quality of life. The functioning's on which human flourishing depends include such elementary things as being alive, being well-nourished and in good health, moving about freely, and so on. It can also include more complex functioning's, such as having self-respect and respect of others, and taking part in the life of the community (including 'appearing in public without shame'), on which Adam Smith in particular presented an extraordinarily insightful analysis in his *Wealth of Nations* [7, p. 357][9].

Therefore, the combination of the different "functioning's", i.e., of different things to do, and/or of different ways of behaving and of being, constitutes the focus of person's life, and the extent of their concrete achievements represent the degree of each person *self-achievement.* On the other hand, *person's capabilities* consist

> in the set of combinations of functioning's from which the person can choose any one combination. Thus, the "capability set" stands for *the actual freedom of choice* a person has over the alternative lives that he or she can lead [7, p. 357].

To sum up, the "basal space" of Sen's theory of justice is characteristically constituted by what "people are able to be or are able to do", rather than in terms of the means or resources they possess. In this way, for the construction of a suitable "social choice theory" based on these *personalistic* principles,

> individual claims are to be assessed neither just by the incomes, resources or primary goods people respectively have, nor only with reference to the pleasures or utilities they enjoy, but in terms of *the freedom they have to choose between different ways of living* they can have reason to value [7, pp. 357-358].

Then, in a social choice theory, of which information basal space is given in terms of "capabilities and functionings", the following question becomes fundamental. Which is the *aggregation principle* of the different variables and of the different interpersonal comparisons, able to take note of variations of individual conditions and circumstances, in which even the underlying personal values and preferences play an essential role?

As we anticipated, as far as we introduce in Rawls' maximin aggregation principle the two axioms of "extended identity" among different welfare rankings, this principle can be extended as well to Sen's SCF theory, as far as we can suppose a suitable real-valued quantitative parametrization of the welfare aggregates involved (see [7, pp. 210-220] for a formal treatment). Indeed, by these axioms we can formalize, on the one side, the ethical principle of the interpersonal/intercultural dialogue as "extended sympathy" among different persons – that is, the ability of posing him/herself in the situation of the other one, "taking on" also his/her personal and ethical preferences and tastes. On the other side, by the same axioms, we can formalize the formation/comparability of homogenous groups of persons in a heterogeneous society, because sharing the same set of personal conditions/preferences. In both case, indeed, the formal condition consists in the possibility that there exists in our modeling an one-to-one correspondence from the set of persons H to H itself, such that (the objective/subjective situation of) a

---

[9] Apart from the reference to Smith's masterpiece, Sen quotes in note another less famous book of A. Smith, *The Theory of Moral Sentiments* [51, pp. 351-352], of which Sen himself cured a re-edition with his own Preface.

person j can be mapped onto (the objective/subjective situation of) a person k, that is, $k = \rho(j)$, despite "originally" $k \neq j$.

Of course, Sen's theory of SCF depending on a theory of justice as equity whose basal information space is given in terms of capabilities and functioning's poses significant logical and then computational problems for their modeling.

4 Conclusion: the computational challenge of Sen's theory

4.1 The computational challenge: the dynamic weighing of variables in Sen's theory

Sen well synthesizes the deep problems related with variable weighing in statistical modeling of his SCF theory in the following passage of his book [7, pp. 358-359].

> The heterogeneity of the components in the basal space, such as different functioning's, points inevitably to the need to weigh them against one another. This applies to all approaches that respect plurality in one form or another, including the Rawlsian focus on primary goods, or the Aristotelian focus on functioning's and capabilities (which is present also in other theories that take note of different aspects of the quality of life).

For instance, we can "close our eyes" by simply *assuming* that there exists something homogenous like "the income" by which everyone's advantages can be judged and impersonally compared, but this cannot solve the problem rather it avoids it. Indeed,

> real income comparisons involve aggregation over different commodities, and in judging comparative individual advantages there is the further problem of interpersonal comparisons, taking note of variations of individual conditions and circumstances. It is, of course, possible to reflect these variations in values of 'adjusted income' that can be appropriately defined, but that is only another way of stating the same problem, requiring that attention be paid to the valuation of heterogeneous factors, though expressed in the 'indirect' space of equivalent incomes. Measurements in the direct space (e.g. quality of life, or capability indicators) and those in the indirect space (e.g. equivalent incomes) would have a tight correspondence with each other, related to the underlying values on which both the normative exercises are based. *One way or another, the issue of valuation and weighting has to be faced.*

Of course, all this means that the usual mathematical and computational tools of statistical analysis used in social and economic sciences appear inadequate, as far as there are based on the necessity of supposing a *fixed and then unique set of weights* for evaluating the variables. Indeed, the problem depends on the high variability of the circumstances that make in principle impossible to arriving at the definition of a unique and fixed set of weights, because of the consequent ever changing character of the inner higher order correlations in the data set. The same problem that computer scientists are faced when dealing with *data streams.* It is therefore not casual that the main criticism to Sen's SCF theory is that

> it is not practically useable because of the weighing problem. In fact, however, diversity of ingredients is only an invitation *to address the problem of relative values and weights,* not an admonition to resign and go home. Indeed, when the valuation of inescapably diverse

> concerns is done implicitly as in the measurement of GDP through the use of market prices for distinct commodities we simply withhold our responsibility to evaluate, going instead for the mechanical use of some relative values that may have been fixed without any relevance to normative assessment (market prices can be *hugely inappropriate for ethical evaluations*, when there are externalities, asymmetric information and big inequalities in income distribution) [7, p. 369].

On the other hand, Sen emphasizes another fundamental factor making simply absurd the argument sometimes proposed that the solution for making usable and even "operational" the capability approach only if it comes with a set of "given weights on different functioning's in some fixed list of relevant capabilities" – like for instance in Martha C. Nussbaum's approach (see ch. 2 of [31]). This factor is strictly related with our communication age in which valuation and weighing is *in real time* systematically influenced by the public discussion and scrutiny on it.

> This need not, however, derail evaluation of public policy, or disable welfare-economic evaluations. The capability approach is entirely consistent with reliance on partial rankings and limited agreements. *The main task is to get the weights – or ranges of weights – appropriate for the comparative judgements that can be reached through reasoning, and if the result is a partial ranking, then we can make precisely those judgements that a partial ranking allows*. There is no obligation to feel compelled to opine on every comparative assessment that can be proposed [7, p. 369] (Italics are ours).

4.2 The computational challenge: the necessity of AI support and the issue of dynamic weighing in machine learning

It is precisely at this point that the relevance of the AI support to social choice theory becomes evident. Indeed, if the problem mainly consists in getting "the weights appropriate for the comparative judgements that can be reached through reasoning", because of the complexity of these judgements and the necessity to be continuously updated often in real time, the support of the AI to human reasoning becomes essential. On this regard, for concluding this Section, it is essential to discuss briefly the core of the actual discussion about the different *machine learning*[10]models in AI systems because their

---

[10] By "machine learning" computer scientists generally intend the ability of a machine of simulating human brain ability of "learning" from its "past experience" without any further intervention of the programmer. Namely, in classical or *symbolic* AI computational systems such as those based on M. Minsky's *frame theory* [55] – which is at the basis of contemporary "object-oriented programming" technique –, when the machine encounters a new situation it selects from memory a given data-structure or *frame* for representing a stereotyped situation, whose details can be adapted to fit reality. Effectively, a frame is a network of nodes and relations, whose top levels are fixed because representing what is supposed to be "always true" (effectively, "true" with the highest probability) in a given situation, and lower levels have many terminals or "slots" to be filled with specific instances or data, according to specified conditions its assignments must meet. Of course, "frames" are *programmed* data-structures. Machine learning using ANN deep learning wants precisely to avoid this human programmer intervention in the definition of the memorized reference patterns. In this sense, it is at the basis of the so-called *non-symbolic* approach to machine leaning, also because in the case of the "big data" this human programmer intervention is effectively impossible, because

strict relationship with the problem of the statistical variable weighing that emerges therefore as the true challenge both for the natural and the artificial reasoning in any realm of contemporary sciences.

Effectively, the actual state of art in machine learning in AI systems is strictly related with the different *deep learning* algorithms inspired by the so-called *artificial neural networks* (ANN) approach in the so-called *non-symbolic* AI, necessary for dealing with large sets of statistical data and of their inner and variable correlations (see note 10), as it is the case of our social choice theories.

Without entering here into the technicalities of ANN approach in machine learning, and despite the demonstrated utility of the different and more evolved learning algorithms in solving specific problems with "big data" (see [34], for an updated review), it is significant for us for understanding why the actual "deep learning" approach in AI-decisional support for human reasoning is certainly useful but not yet sufficient *in principle* for addressing to the problems emphasized by Sen of a true *dynamic* weighing of the variables in social choice theory – and not only in it.

Indeed, all the actual "deep learning" models in AI systems are based on a *statistical* (not *dynamic*) approach to machine learning. This is essentially a modeling in which there exists a sharp division between the "learning phase" onto some representative sample of the database for *fixing* the proper set of weights representing the *higher order correlations* among the variables in the dataset, and the "testing" phase for applying fruitfully the learned weights configuration onto the whole database. Therefore, it is evident that this approach cannot work when we have to deal in machine learning not only with huge, even though *static* bases of data, but with *dynamic* streams of data, in which there exists an indefinite and unpredictable – because continuously changing – number of possible correlations among an anyway finite set of data. That is, we are faced again with the *infinity* challenge in computational systems. This means that such an approach cannot work in the actual social and economic contexts, in which the fast information exchange among the social and economic actors distributed all-over the world determines the sudden formation/destruction of *long-range correlations* among their behaviors. In such cases, like, on the other side, in Sen's modeling of SCFs in which economic and social actors are directly involved in the process to be modeled, the weights cannot be fixed in any way, as Sen emphasized many times. In other terms, in these cases we need ANNs in *continuous learning,* just it happens in biological neural networks in which no absolute separation between the learning and testing phases occurs. The brain neurons, constituting the hidden layers of biological NNs, indeed, not only are always *oscillating*, otherwise their EEG would become flat, but the EEG traces in healthy brains are characterized by the highest variability, without any recurrent periodicity, because a living brain is physically a *dissipative* or *open system* in continuous exchange with its environment.

Unfortunately, for the actual statistical approach to ANNs that are based on statistical mechanics and then *must* model ANNs as *closed* systems non interacting with their environments, "continuous learning" means that the non-linear behavior of the net

---

of the complexity of the data structures involved, characterized by millions and ever billions of parameters to be taken into account.

becomes immediately *chaotic* and then in principle intractable for whichever statistical learning algorithm such as, for instance, the well-known *gradient descent* algorithm that, with several updating and corrections, is the actual standard paradigm in AI machine learning, since 80's of the last century [34, 35, 36].

According to this approach, the deep learning in the many layers of the hidden neurons between the input and the output layers of the net, and whose dynamics is governed by a non-linear *sigmoid* function σ[11], consists in an optimization process by the stochastic *gradient descent algorithm*. That is, it consists in the calculus of the minimum of an error-function ("loss function"), reducing recursively the distance between the actual and the desired output of the net. This happens by *back-propagating* recursively the error for redefining the statistical weight configuration of the connections between the input and the hidden neurons of the net, till the *absolute minimum* of error function is reached [36] (see **Fig. 1**). This means that the back-propagation learning algorithm is a *supervised learning algorithm* since the optimization process supposes the existence of a learning target to be reached, depending on a "teacher", effectively the human programmer.

The separation between the learning and the testing phases in this *statistical* standard paradigm of ANNs depends then theoretically on the necessity of defining through the learning phase one only fixed set of neuron weights, and therefore one only "energy landscape" for the dynamics of the net, because determining one only "potential

[11] Effectively, the non-linear *sigmoid function*: $\sigma(x) = \frac{1}{1+e^{-x}}$, and overall its immediate relative the *hyperbolic tangent function* $\tanh x = \left(\frac{e^x - e^{-x}}{e^x + e^{-x}}\right) = \frac{\sinh x}{\cosh x}$ are what characterizes the learning algorithms of the actual ANNs, instead of the *linear* stepwise 1/0 activation function of Rosenblatt's early *linear* perceptron [53]. This is, effectively, the *Heaviside function:* $H(x) = \frac{d}{dx}\max\{x, 0\}$ for $x \neq 0$, whose value is "zero" for negative arguments and "one" for positive arguments, so to make linear – namely, a summation of the neuron weights each representing a variable of the dataset – the neuron activation function. To sum up, by using the sigmoid as the activation function, the overall neuron output is no longer 0/1, but any real value between 0 and 1, so to allow to calculate the probability that some event (output) happens, given some conditions (input), instead of the simple *no/yes* answer of linear ANNs. The non-linearity is then fundamental for calculating *higher-order correlations* among data. Indeed, mathematically, the Taylor series expansion of *tanh* includes several higher order correlations, much more than *sinh* and *cosh,* of which *tanh* is the ratio, so to answer in principle to the main criticism made by M. Minsky and S. Papert at MIT [54] to early *linear* ANNs structures, such as McCulloch's and Pitt's neural nets [52] and Rosenblatt's linear "perceptron" [53], who first demonstrated that a neural net is able to calculate in a desirable parallel way any Boolean logical function like a Turing Machine. Effectively, because of the prestige of M. Minsky in US computer science community, his criticism was able to block any further research in ANNs, from 60's to 80's of the last century. In 80's a non-linear multilayer perceptron with a sigmoid as activation function for the hidden or "deep" layer neurons [36], and then characterized by the "gradient descent" learning algorithm [35], answered Minsky's main criticism to early linear ANNs. However, only during the last twenty years, with the large availability of ever more powerful processors – essentially arrays of GPUs – for reckoning with the computational weight of this type of ANN architectures we have the actual development of AI-systems based on "deep" machine learning algorithms [34].

function" of the error to be minimized. This, therefore, requires that the system be conceived as *stable at equilibrium,* and then as "isolated" from any other further information exchange with its environment, otherwise the dynamics would lose its *structural stability*. Namely, its weights would be continuously changing like its overall error "energy landscape". Now, this *non-stationary* character of a dynamics is precisely what characterizes the notion of *deterministic chaos* in the theory of non-linear dynamic systems,

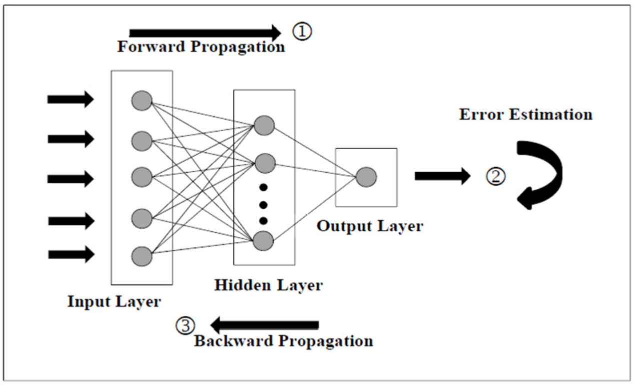

Fig. 1. Elementary scheme of the *back-propagation* ANN architecture with one only hidden layer of neurons endowed with the *gradient-descent* learning algorithm (see **[35, 36]** and note 11 for further explanations).

ANNs included [37, 38][12]. This finally explains because these statistical models of machine learning might work well with "big data" in *static* bases of data, but not in *dynamic* streaming of data.

Moreover, the all-with-all connectivity between input and hidden neurons (see **Fig. 1** above) explains, on the one side, why this computational architecture effectively does not answer – theoretically – the other main objection of Minsky against ANNs. Namely that they are not properly *parallel computation* architectures (for being so, different subsets of the hidden neurons would have to be linked only with subsets of the input neurons and not all-with-all). On the other side, this *total connectivity* explains why this type of machine learning algorithm are so "computational power consuming". Indeed, with the full connectivity, the number of the possibilities to be computed grows *exponentially* with $n!$, where n is the number of neurons of the net. This explains why this machine learning statistical approach spread effectively only during the last twenty years, when computationally powerful GPUs became available at chip prices on the

[12] Effectively, a chaotic trajectory in the phase space of the system is not a purely *random* walk like in stochastic systems. Indeed, it is characterized by unpredictable "jumps" of the dynamics onto a *finite* number of different cycles, with different *timings,* never closed on themselves (i.e., they are not *periodic* so that a chaotic trajectory never passes again where it is already passed), in different *spatial* locations. W. Freeman, uses on this regard an effective cinematographic metaphor: it is like the different frame sequences of a movie, simultaneously projected onto different screens.

market. Even though this approach became known to the wide public only few years ago, effectively after the successful application by Geoffrey E. Hinton in 2012 [39] to the *ImageNet* database of millions of images. An application that earned him the front page on the *New York Times*[13]. On the other side, this "computational weight" of deep learning algorithms explains why the future of machine learning in AI is strictly related with the development of much more computationally powerful *quantum computers*.

It is therefore highly significant that the dynamics of the *biological neural nets* of our *living* brains is *chaotic* [40]. Living brains are evidently open systems in continuous exchange with their environments – otherwise they are dead brains in containers fulfilled with formalin of some neuroanatomy lab.

However, if this chaotic character make in principle despairing any attempt of a *statistical* approach to NNs in continuous learning, this opens the way to a *dynamic* modeling of this type of "deep learning", effectively already developed for biological NNs, but that can be applied also to ANNs. Indeed, the connection with our previous discussion is that such a chaotic dynamics is the demonstrated *macroscopic* manifestation of an underlying *microscopic* dynamics of the *many-body field theory* of dissipative quantum systems, characterizing the fundamental physics of the condensed matter, biological systems included [41, 42].

On the other hand, the reference to dissipative QFT of many-body interacting systems we anticipated many times in this paper as the proper physical paradigm to which also Sen's social theory of the many interacting groups in society must refer, in the case of biological NNs, has another proper justification. It is, indeed, the only possible explanation of the *long-range correlation waves* proper of the oscillatory behavior of mammalian neocortex, as well as of the mammalian thalamocortical interactions during the *action-perception* cycles characterizing their *intentional behavior.* This, on the one hand, suggests an intriguing hypothesis about a *dynamic* mechanism of *multilayered long-term memory* constitution (*dynamic* "deep learning" in quantum computing) in the mammalian brains, using the powerful QFT principle of the *QV-foliation* [41, 42]. On the other hand, the operator algebra formalism underlying this model [43] suggests the development of a new architecture of quantum optical computer based on dissipative QFT in physics, and on Category Theory in logic [11, 44], by which approaching from a new *dynamical* and no longer *statistical* point of view the challenges of an unsupervised machine learning in AI [45], for dealing with data streaming and with dynamic bases of data, both characterized by the sudden formation-dissolution of *long-range correlations* (field phase coherences) among data.

Effectively, the same G. Hinton more recently stated that the future of machine learning in AI is related with the development of *unsupervised* learning algorithms in ANNs. In an interview with Steve Levin on September 15th, 2017 the interviewer so summarized Hinton's new position:

> In back-propagation *labels* or "weights" are used to represent a photo or voice within a brain-like neural layer. The weights are then adjusted and readjusted, layer by layer, until the network can perform an intelligent function with the fewest possible errors.

---

[13] See, https://www.nytimes.com/2012/11/24/science/scientists-see-advances-in-deep-learning-a-part-of-artificial-intelligence.html

> But Hinton suggested that, to get to where neural networks are able to become intelligent on their own, what is known as *unsupervised learning* "I suspect that means getting rid of back-propagation". "I don't think it's how the brain works" he said. "We clearly don't need all the *labeled* data"[14].

This reference of G. Hinton to the computational notion of *labeling* is particularly precious for understanding how "brain works", according to the QFT modeling we illustrated in many papers [11, 42, 41, 46]. Indeed, the notion of "label" pertains to the elementary notions of theoretical computer science, where any *automaton,* Turing Machine (TM) included, can be modeled as a "*labeled* state transition system (LTS)". The different models in AI machine learning can be then usefully characterized according to the different interpretations of the notion of "label".

1. In the early classical or *symbolic* AI machine learning models, such as Minsky's *frame theory* (see above note 10), the label is *an instruction of the program*, more precisely a function of the logical Boolean calculus determining a computational step or "state transition" in the machine, according to the TM symbolic paradigm in computer science.
2. In the *non-symbolic* but *statistical* AI machine learning of the ANNs, the label is a given *configuration of the weights* of the transition probability matrix representing a specific ANN architecture, according to different neuron activation functions. In the case of *supervised* models such as "back-propagation" and similar, for configuring the weights they use the "gradient descent algorithm". As we know, these machine learning models, representing the actual state of the art in AI, are "supervised" because, even though the weight configuration does not depend directly on the programmer like in the "statistical TMs", nevertheless the target representation for calculating the right weight configuration by "back-propagating the error" is put by hand, i.e., it supposes "a teacher". Of course, it is possible to think also at *statistical unsupervised models* of ANNs, which are at the moment purely speculative because generally lacking a suitable *selection criterion* among weight configurations (see however note 15 below).
3. In the *non-symbolic* but *(thermo-)dynamic* quantum models of AI machine learning, the label is a given *condensate of NG-bosons,* each identifying univocally a *long-range correlation* among oscillating neural fields and their (inner/outer) environments, both in biological, and in artificial NNs [42, 11, 41]. This "dynamic labeling" is observer-independent, i.e., *absolutely unsupervised,* and act as a *control parameter* of a given *phase coherence domain,* constituting physically in QFT or socially in Sen's theory (in both cases, coalgebraically) the logical reference (coalgebraic semantics, including the system and its physical/social environment) making true the relative proposition of a Boolean algebra calculus. This class of biological and artificial NNs represent therefore a QFT instantiation of a "*thermodynamic* LTS", where effectively the "states" are different *phases* of matter in physical, biological and social natural realms, and/or different *controlled phases* of interfering coherent light waves (photon condensates) in the two branches (the principal and the reference one) of an interferometer, in our quantum optics implementation of the same principle [11]. In this case, because we

[14] For a more articulated and updated criticism to the supervised machine learning in AI, see Gary Marcus' review paper [56].

work with "open" systems, in far-from-equilibrium conditions and therefore we are outside Gibbs' statistical mechanics paradigm, differently from QM, we can pass through different entangled phases (states), using each time *the minimum free energy function* as a *generalized (universal) dynamic selection criterion* among states (phases), linking systematically free energy dissipation (entropy) with information[15]. This is the core of the *doubling of the degrees of freedom* in QFT we illustrated elsewhere [42, 11, 41]. See on this regard Eq. (7) in [11], where it is made explicit the dynamic labeling $\mathcal{N}$ of the NG-boson condensates in the state (phase) transition rule (i.e., the "doubled" Bogoliubov transform) governing the machine learning of our QFT implementation of a LTS, and where the tanh plays not casually an essential role (see note 11 above).

Finally, for concluding this contribution, we want to recall the continuity of this thermodynamic interpretation of machine learning with Lorenzo Magnani's recent characterization of deep learning processes in humans, with respect to the actual AI systems – and specifically with the "AlphaGo" program, taken as a testbed example [47]. This specificity consists in the continuous interaction of humans with their environments, defined as the *eco-cognitive openness* or the *unlocked* character of human creative abductive processes, with respect to the *locked* character of the AlphaGo abductive strategies, and generally of the AI systems based onto statistical deep-learning algorithms. The dissipative QFT interpretation of mammalian and human neural dynamics [42, 41], and the consequent quantum logical calculus on a coalgebraic basis we can implement in it [11] might be then considered as a modeling of Magnani's "eco-cognitive" abductive processes also because they share the same *semiotic* and then algebraic background in logic [15].

References

1. Hansson, S. O., Hendricks, V. F. (eds.): Introduction to Formal Philosophy, Springer, Berlin-New York (2018).
2. Endriss, U.: Logic and social choice theory. In: Gupta, A., Van Benthem J. (eds.) Logic and Philosophy Today, pp. 333-377. College Pubblications, London (2011).

[15] At this point, it is significant to emphasize that in US, Todd L. Hylton and his collaborators recently launched the *Thermodynamic Computer Project* (see https://knowm.org/thermodynamic-computing/). For instance, the so-called "thermodynamic NNs" are trying to approach a type of unsupervised solution in machine learning using the minimum free energy (in Helmholtz' formalization within the statistical mechanics approach) as a selection optimization criterion. They remain, however, in the paradigm of the statistical interpretation of thermodynamics, also for modeling neurons interacting with (their representation of the) environment, using an ingenious, but anthropomorphic logic of "question/answer" [57]. In this way, the dynamics of such a "thermodynamic neural network" is a sort of stroboscopic approximation (i.e., limited to one phase or to a small number of close phases) of what in QFT is a continuous process of phase transitions (macroscopically, a chaotic dynamics) and then, when implemented computationally, a *universal* thermodynamic LTS. In this limited sense, this class of ANNs is straddling (2) and (3) types of unsupervised machine learning algorithms.

3. Bergson, A.: A reformulation of certain aspects of welfare economics. Quarterly Journal of Economics 52(2), 310-334 (1938).
4. Samuelson, P. A.: Foundations of economic analysis. Harvard UP, Cambridge, MA (1947).
5. Arrow, K. J.: Social choice and individual values. 2nd edn. Yale UP, New Haven & London (1963).
6. Sen A. K.: Collective Choice and Social Welfare. Elsevier, Amsterdam (1979).
7. Sen A. K.: Collective Choice and Social Welfare. Expanded Edition. Penguin Ltd. Kindle Edition, London (2017).
8. Sen A. K.: The impossibility of a Paretian liberal. Journal of Political Economy 78(1) 152-157 (1970).
9. Rawls, J.: A Theory of Justice. Oxford UP, Oxford, UK (1971).
10. Stiglitz, J. E., Sen A. K., Fitoussi, J.-P.: Mis-meauring our lives. Why GDP doesn't add up. The report by the Commission on the measurement of economic performance and social progress. With a foreword by Nicolas Sarkozy. New Press, New York (2010).
11. Basti, G., Capolupo, A., Vitiello G.: Quantum Field Theory and Coalgebraic Logic in Theoretical Computer Science. Prog. in Bioph. & Mol. Biol. Special Issue: Quantum information models in biology, from molecular biology to cognition. 130, 39-52 (2017).
12. Sen, A. K.: The Idea of Justice. Penguin Book, London (2010).
13. Sen, A. K.: Development as freedom. Anchor Books, New York (2010).
14. Tirole, J.: Economics for the common good. Princeton, NJ, Princeton UP (2017).
15. Basti, G.: The Post-Modern Transcendental of Language in Science and Philosophy. In: Delic, Z. (ed.) Epistemology and Transformation of Knowledge in a Global Age, pp. 35-62, InTech, London (2017).
16. Mulgan, G.: Big Mind. How collective intelligence can change our world, Princeton UP, Princeton, NJ (2018).
17. Kirman, A., Sonderman, D.: Arrow's Theorem, Many Agents, and Invisible Dictators. Journal of Economical Theory 5(2), 267-277 (1972).
18. Mihara, H. R.: Arrow's theorem and Turing computability. Economic Theory 10(2) 257-276 (1997).
19. Sen, A. K.: The Informational Basis of Social Choice. In: Maskin, E., Sen, A. K. (eds.) The Arrow Impossibility Theorem, pp. 67-100, Columbia UP, New York, (2014).
20. Rawls, J.: Jusitice as Fairness. The Philosophical Review 67(2), 164-194 (1958).
21. Rawls, J.: Justice as fairness: political not metaphysical. Philosophy and Public Affairs 14, 223-252 (1985).
22. Rawls, J.: Political Liberalism. Columbia UP, New York (1993).
23. Rawls, J.: Justice as Farness. A Restatement. Kelly, E. (ed.) The Belknap Press of Harvard UP, Cambridge, MA & London UK (2001).
24. Kant, I.: Grundlegung zur Metaphysik der Sitten. Johann Friedrich Hartnock, Riga (1785).
25. Sidgwick, H.: The Methods of Ethics, 3° vol. McMillan, London (1874).
26. Hare, R. M.: Freedom and Reason. Oxford UP, Oxford, UK (1963).
27. Hare, R. M.: The Language of Morals. Oxford UP, Oxford, UK (1963).
28. Rawls, J.: The priority of rights and the idea of good. Philosohy and Public Affairs 17(4), 251-276 (1988).
29. Suppes, P.: Some Formal Models of Grading Principles. Synthese 16(3-4) 284–306 (1966).
30. Nussbaum, M. C.: Human Functioning and Social Justice: In Defense of Aristotelian Essentialism. Political Theory 20(2), 202-246 (1992).

31. Nussbaum, M. C.: Creating Capabilities. Harvard UP, Cambridge, MA (2011).
32. Alkire, S.: Valuing Freedoms: Sen’s Capability Approach and Poverty Reduction, Oxford UP, New York (2002).
33. Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M., Ballon P.: Multidimensional Poverty Measurement and Analysis. Oxford UP, Oxford, UK (2015).
34. Nwankpa, C. E., Ijomah, W., Gachagan A., Marshall, S.: Activation Functions: Comparison of Trends in Practice and Research for Deep Learning, 8 November 2018. https://arXiv.org/arXiv:1811.03378v1, last accessed 30 June 2019.
35. Werbos, P.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD thesis. Harvard UP, Boston, MA (1974).
36. Rumelhart, D. E., Hinton, G. E, Williams R. J.: Learning representations by back-propagating errors. Nature 323(6088) 533-536 (1986).
37. Hilborn, R.: Chaos and non-linear dynamics. Oxford UP, Oxford (1994).
38. Viana, M.: What's new on Lorenz strange attractors. Math. Intelligencer 22(3) 6-19 (2000).
39. Krizhevsky, A., Sutskever, I., Hinton, G. E.: ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, pp. 1097-1105. Curran Associates Inc., Lake Tahoe, Nevada, (2012).
40. Freeman, W. J.: How brains make up their minds. Columbia UP, New York (2001).
41. Freeman, W. J., Vitiello, G.: Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Physics of Life Reviews 3(2) 93-118 (2006).
42. Freeman, W. J., Vitiello, G.: Dissipation and spontaneous symmetry breaking in brain dynamics. Journal of Physics A: Mathematical and Theoretical 41(30) 304042 (2008).
43. Blasone, M., Jizba, P., Vitiello G., Quantum field theory and its macroscopic manifestations. Boson condensation, ordered patterns and topological defects. Imperial College Press, London (2011).
44. Rasetti, M., Marletto, P.: Quantum physics, topology, formal languages, computation: a categorial view - an homage to David Hilbert. Perspectives on Science 22(1) 98-114 (2014).
45. Rasetti, M., Merelli, E.: Topological field theory of data: mining data beyond complex networks. In: Contucci, P., Giardinà, C. (eds.) Advances in disordered systems, random processes and some applications, pp. 1-42. Cambridge, UP, Cambridge UK (2017).
46. Freeman, W., Vitiello, G.: Matter and mind are entangled in two streams of images guiding behavior and informing the subject through awareness. Mind & Matter 14(1) 7-24 (2016).
47. Magnani, L.: AlphaGo, Locked Strategies, and Eco-Cognitive Openness. Philosophies 4 8-24 (2019).
48. Condorcet, M. de : Essai sur l’application de l’analyse à la probabilité des decisions rendues à la pluralité des voix. L'Imprimerie Royale, Paris (1785).
49. Kant, I.: Fundamental Principles of the Metaphysics of Ethics. Engl. Transl. by T. K. Abbott. Longmans, London (1907).
50. Pattanaik, P. K.: Social choice and voting. In: Hansson, S. O., Hendricks, V. F. (eds.) Introduction to Formal Philosophy, pp. 839-853. Springer, Berlin-New York (2018).
51. Smith, A.: The Theory of Moral Sentiments, 1776, Sen, A. K. (ed.). Penguin, London (2009).

52. McCulloch, W. S., Pitts, W. H.: A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics 5(4) 115-133 (1943).
53. Rosenblatt, F.: Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms. Cornell UP, Buffalo, NY (1961).
54. Minsky M., Papert, S. : Perceptrons. An introduction to computational geometry. Second Edition. MIT Press, Cambridge, MA (1987).
55. Minsky, M.: Frame theory. In: Johnson-Laird, P. N., Wason, P.C. (eds.) Thinking: Reasings in Cognitive Science, pp. 355-376. Cambridge, UP, Cambridge MA (1977),
56. Marcus, G.: Deep learning: a critical appraisal. https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf. Accessed 15 July 2019.
57. Fry, R. L.: Physical intelligence and thermodynamic computing. Entropy 19, 107, (2017).