Advances in Complex Systems
Vol. 23, No. 7 (2020) 2020001 (7 pages)

DOI: 10.1142/S0219525920200015

DESIGNING SYSTEMS BOTTOM UP: FACETS AND PROBLEMS

FRANK SCHWEITZER
Chair of Systems Design,
ETH Zurich, Switzerland
acs@wspc.com

Received 21 May 2021
Accepted 25 May 2021
Published 26 June 2021

Systems design utilizes top-down and bottom-up approaches to influence social or economic systems such that a desired outcome is obtained. We characterize different approaches like network controllability, network interventions, nudging and mechanism design and discuss the problems involved. We argue that systems design cannot be reduced to solving complex optimization problems.

Keywords: Mechanism design; social planner; nudging; network controllability; interventions.

These days, the Corona crisis provides us with ample opportunities to watch systems design at work. Social planners in all countries have clear-cut ideas of desired outcomes that should be achieved by various types of systemic interventions. Their tool box has impressively expanded. Classical methods for steering the outcome by means of legal restrictions, punishments and incentives are now combined with biased information campaigns, rule changes, legislative empowerment and social pressure. Thus it is timely to have a look at systems design from a broader perspective.

Top-down versus bottom-up. Systems design, in general terms, denotes an approach to manipulate technical, environmental, social or economic systems such that a desired outcome is obtained. This can be achieved in a bottom-up or a top-down manner or a combination of both.[a] The top-down method focuses on the boundary conditions under which systems evolve. In socio-economic systems one tries to adjust global variables, for example tax rates, customs tariffs or legal

[a] See: Schweitzer, F., The Bigger Picture: Complexity Meets Systems Design, in *Design. Tales of Science and Innovation*, Folkers, G. and Schmid, M. (eds.) (Zürich: Chronos, 2019), pp. 77–86, https://www.sg.ethz.ch/publications/2019/schweitzer2019the-bigger-picture/

frameworks, to obtain the desired outcome. The focus is on the *macro perspective*, as in macro economics or macro sociology.

In contrast, the *bottom-up approach* focuses on the system elements, commonly referred to as agents. These can be influenced in two ways, either through their internal dynamics or through their interactions with other agents. Hence, the bottom up method takes a *micro perspective*. The desired systemic behavior then manifests itself as an *emergent property* resulting from the interactions of a large number of agents.

Given that socio-economic systems are inherently *complex systems* comprising a large number of heterogeneous entities, the top-down approach is rather limited in controlling the outcome of collective interactions. This makes bottom-up methods interesting alternatives.

Mechanism design. From the perspective of the social planner, a system can be possibly improved by providing a balanced solution that nobody has an incentive to change. Game theory, for example, provides concepts of Pareto optimality or Nash equilibria, to describe the stability and the efficiency, i.e., the aggregated welfare, of solutions. This approach transforms the problem of systems design into a *optimization* problem. It requires (i) to know the optimization function, (ii) to ensure that it can be solved and (iii) to implement the optimal solution. None of these requirements are usually satisfied or attainable for practical purposes. Therefore, this approach relies on formal methods involving proofs about the existence of optimal solutions. It is restricted to a small class of problems which are analytically tractable.

Mechanism design, an established research field of micro economics, is such an approach. Instead of studying the emergent properties of multi-agent interactions, mechanism design aims at restricting these interactions by design. Precisely, a *rule* defines constraints for interactions. Such rules differ from, e.g., traffic rules to drive on the left/right side, which are not "designed" but *emerge* from the collective interactions. Textbook examples for mechanism design are auction rules. In auctions agents have to bear the consequences of their decisions: if they bid too high, they may win, but have to pay a price too high. Thus, rational agents would not offer more than the assumed value in such auctions. To enforce that, on the other hand, agents will not understate, the Vickrey rule determines that winning agents will only pay the second highest price. Hence, defining such rules appropriately ensures a fair outcome, while maximizing the resulting gain.

Auction design aims to *reveal information* about the true preferences of agents, which would not be revealed otherwise. But in most cases one is rather interested in *changing* such preferences, for instance by suppressing or promoting alternatives. Market design, another special application of mechanism design, tackles this problem from an optimization perspective. It provides different algorithms to solve the matching of two sides of a market, for example matching doctors and patients, students and internships, or arranging stable marriages. Again, the main idea is to choose appropriate rules that restrict certain "interactions". There are possible relationships that should not establish, while others with a lower likelihood should be favored, in order to optimize the outcome on the systemic level. Ideally, the design ensures that no agent has an incentive to deviate from the proposed solution.

From solutions to problems. What are the problems involved in mechanism design? Conceptually, it requires a social planner with full control of the system. Agents can only announce their preferences, but they do not make decisions. In fact, agents do not interact directly. They only communicate with a central authority that solves an optimization problem for them, algorithmically. Hence, with these restrictions provided by a social planner we can argue that mechanism design should be seen as a top-down rather than a bottom-up approach.

One question is how many real-world problems *can* be *solved* this way. In most cases, the underlying optimization function is not known or too complicated to be solved. More importantly, *should* real-world problems also be solved this way? Recent debates about ethical implications from algorithmic decision-making, for instance in hospital triage systems to determine the priority of patients' treatments, indicate a rather critical stance.

The usual counterargument emphasizes that mechanism design ensures at least an optimal outcome that nobody has an interest to deviate from, which would not be reached with unrestricted interactions. But that holds, in the best case, only in theory. In reality, there are ample ways of market manipulation, ranging from corruption to fraudulent actions and state interventions. Existing theories of mechanism design do not adequately reflect these manipulation possibilities.

Coping with evolution. In addition to methodological issues, there are fundamental issues. The fact that an optimal solution exists does not imply that it is attainable in practice. This is not the result of our inabilities, it is the consequence of the structure of real-world problems. Solutions need to be implemented, which requires time and resources. We never start from scratch, but build on existing structures. Therefore, *path dependence* in the dynamics of socio-economic systems often makes it impossible to really change direction to reach the desired state.

Further, while implementing a seemingly optimal solution, the system continues to evolve. There is a *co-evolution* of systemic changes induced by agents and their possible actions. *Self-organization*, i.e., the ability of complex systems to generate collective states with new properties, will not always result in desired outcomes. As often, the result is an inefficient, unwanted system state, examples ranging from stock market crashes to failed political states. These shortcomings have lead to the concept of guided self-organization, to avoid unwanted outcomes by additional bottom-up or top-down control of the system dynamics. But socio-economic systems are adaptive, i.e., they respond to internal or external changes. Hence, we deal with moving targets rather than static goals.

Specifying the goal. Most concepts of systems design require to specify the desired systemic state. This means appropriate measures must be devised to capture the system state. This in turn creates a problem in itself. Most often, we can only compare system states relative to another rather than measuring how close or far we are with respect to a certain optimum.

Economic key variables only provide us with aggregated *numbers*. Maximizing the GDP, for example, does not tell anything about the optimality of the outcome with respect to other problem dimensions, for instance the resilience of the system or the fair distribution of income. In theory, such additional requirements can be included as constraints for the optimization, but in practice the feasibility is limited. Therefore an optimally designed solution from the social planner may not be satisfactory with respect to other criteria or special interests of individuals or groups of agents. Statistical physics has developed the telling notion of *frustrated systems* to express such situations.

Considering all these arguments, it becomes apparent that systems design cannot be reduced to solving complex optimization problems, for principal reasons. The social planner may have the legal rights to determine optimization functions, rules or incentives, to control a system. But all this does not ensure that the desired outcome can be indeed obtained, nor that it is still optimal when it is eventually reached.

Alternative bottom-up approaches. What is left for us? We keep the main goal to influence systems such that a desired outcome is obtained. But we reject the idea that this goal can be reached by proofs about the existence of such desired states, or by solving optimization problems. Instead, we have to test under which conditions systems are enabled to find suitable states *on their own*, without a social planner specifying the outcome. In essence, this means to give degrees of freedom, for instance regarding the interaction of agents, back to the *system*, while mechanism design has reduced these degrees of freedom even further.

This has two modeling implications. The first one: Respect the *system*. Everyone knows that "the whole is greater than the sum of it parts" (Aristotle) or that "more is different" (Phil Anderson). Still, this is often not reflected in the way system models are built and analyzed. Many theoretical methods assume that collective interactions can be decomposed. Game theoretical models, for example, routinely reduce multi-agent interactions to 2-person games. This has the advantage that both the rules and the resulting payoff are defined, the dynamics can be predicted and it becomes clear how to obtain a desired outcome. But the advantage of calculating possible equilibria is outweighed by the restrictions of generalizing from this very limited setting.

The second implication: Respect the *eigendynamics*. Everyone knows that implementing solutions against systemic forces is much harder than implementing solutions aligned to such forces. But to leverage these forces, we first need to understand the eigendynamics, i.e., the dynamics of the systems before any interventions. This would allow us, for example, to identify those agents with the most impact on the systemic dynamics and to utilize their influence to steer the system. There are different approaches to achieve this, which are shortly discussed in the following.

Network approaches. *Network controllability*[b] is the most general approach in that it studies the *principal ability* to drive a system into *any state* that is compatible with the system's structure and dynamics. It uses methods from control theory to identify *driver nodes*, i.e., agents that can be utilized to steer the system. Only a small subset of agents is targeted to partially or completely control the system. This approach has the advantage to formally identify those sets of agents that are instrumental in influencing the outcome. However, as a disadvantage it requires information about the network topology and the dynamics that runs on this network. Changes in interactions, the evolution of network structures or changes in the dynamics are not considered.

A specific example is the control of reputation spillovers between firms by means of their ownership relations. If firms fail to comply, for instance, to ethical standards of production, their shareholders do not only suffer economically, they also risk their reputation. Thus, reputation can be channeled through the ownership network. But which firms should be targeted to improve the situation? Are the most reputed firms also the most efficient driver nodes? Such questions can be addressed both formally, using the concept of network controllability, and by means of data-driven modeling, building on data about ownership relations.

Network interventions take a broader perspective. One analyzes possible *incentives* that should be provided to the agent to bias its decision or behavior towards a desired systemic outcome. Generally, these incentives influence the *utility function* of an agent, either reducing costs or increasing benefits. Rational agents are assumed to choose outcomes of higher utility. Importantly, because of the *network* structure the impact of such decisions will propagate through the system, this way influencing agents that were not targeted directly. The method uses systemic feedback mechanisms to amplify or suppress certain trends. This approach has the advantage of comparably low costs as it targets only a small number of agents. On the other hand, it requires to identify these agents and the type and amount of interventions, and to forecast the impact on the remaining system.

In the absence of general methods, most often agent-based simulations are used. For example, firms engage in R&D collaborations in order to expand their knowledge stock. If they no longer benefit from the knowledge exchange, they may decide to leave the collaboration network. This impacts other firms which may also leave, this way causing dropout cascades. In simulations one can test various intervention strategies to prevent such cascades, for example by reducing the costs of specific firms, or by the targeted removal of inefficient firms.

[b] See the topical issue on network controllability in *Advances in Complex Systems* (Guest editors: A. Li and Y.-Y. Liu), Vol. 22 (2019), https://www.worldscientific.com/toc/acs/22/07n08.

Soft influence. *Nudging* is a comparably softer way of influencing agents because it provides *information* rather than incentives and leverages psychological mechanisms to steer decisions in certain directions. Agents are presented alternatives in a way that they may choose the one desired by the social planner. Also, social reward mechanisms like reputation scores play a role. At difference with network interventions, agents base their decisions on additional, also on different information. For example, instead of their individual payoff, they may take the total payoff of the population into account because that is presented to them. This neither changes the rules of interactions, nor the payoffs, but may still induce a desired decision.

The advantage, as well as the disadvantage, comes from the subtle way of influencing agents. This type of manipulation is also called *paternalism* in behavioral economics. If the social planner constantly targets agents with a particular information to nudge them in their decisions, they hardly learn about possible other information that would bias them into the opposite direction. This becomes a real problem if social and public media reinforce this preferred information, this way aligning to the implicit goals of the social planner in an nontransparent manner.

Application areas for nudging range from marketing of products in music videos to the dissemination of fake news in social media, the public praise of well-behaved citizen, and political influence on selecting broadcasting news. Thanks to the availability of big data from social media these phenomena can be measured and explained. Modeling such processes, however, is rather difficult as it requires a quantitative understanding of behavioral responses.

Acceptable rather than optimal. Agent-based models are the prime tool for systems design in that they allow to test possible interventions in an exploratory manner. Such bottom-up interventions should consider more than just strategies or utilities or other means of rational behavior. Interestingly, social and psychological mechanisms come into play, for example emotional influence, persuasion, support. Instead of pure enforcement, social feedback mechanisms like the herding effect provide efficient ways to amplify a desired trend. Additionally, it sometimes requires some sort of irrationality, e.g., some level of randomness, to escape from a inferior lock-in state, to reach better ones.

The major question, when *simulating* the system dynamics, is no longer about finding *the* optimal solution. Will agents converge to an *acceptable* solution in a finite time, with finite information? To obtain this goal, the bottom-up approach of systems design can leverage two very effective mechanisms: (i) to influence the information that agents take into account in their decisions and (ii) to establish feedback cycles such that agents are directly confronted with the positive or negative consequences of their decisions.

Information feedback is key. Studies on the wisdom of crowds or on game theoretical problems have demonstrated that *more* information about the decisions of others or about possible payoffs may lead to *worse* outcomes on the systemic level.

On the contrary, more information or stronger social coupling can drive the collective dynamics into suboptimal states. To provide less or only selected information could help to mitigate such situations.

Feedback cycles to relate actions and outcome in a more transparent manner can also build on information. This can be private information to help a single agent to understand the consequences of decisions, but also public information, for example about reputation spillovers. Further, the feedback between actions and consequences can be enhanced through legal regulations. For example, the moral hazard problem where individuals take the benefits of their actions while the general public has to bear the costs and the risks could be tackled by a tighter legal or informational coupling.

To what extent social or economic systems can really be changed by these design methods, remains an open problem. It touches upon the *implementation* of these measures, which points back to the social planner who is in control. But who controls the social planner? This is the question more than ever. With our scientific contributions, we can only inform about possible and impossible routes toward designing systemic outcomes. Whether they are taken up, is beyond our influence.