Multi-Scale Resolution of Human Social Systems: A Synergistic Paradigm for Simulating Minds and Society

Mark G. Orr

March 31, 2018

1 Introduction

Recently, we put forth an initial sketch of what we call the *Resolution Thesis*[3]. The thesis holds that 1) models of cognition will be improved given constraints from the structure and dynamics of the social systems in which they are supposed to be embedded, and 2) the resolution of social simulations of agents will be improved given constraints from cognitive first principles[1]. This thesis reflects a variety of motivations, the most obvious being the observation that there is little overlap between the cognitive sciences and the generative social science approach[*REF EPSTEIN, 2008], both of which rely heavily on computer simulation to understand aspects of human systems, albeit different aspects of human systems with respect to scale. The former focuses almost exclusively on the mind as a scientific object of study for which the lion's share of simulation efforts reflect methods that represent a generalizable conception and the latter emphasizes multiple aspects of social systems, the mind being only one of these aspects. For this latter case, there is little relation to the vast experimental evidence of how the mind operates or what are the central theoretical entities that compose the mind. To a large degree, the resolution thesis is a recognition that an interdependence between cognitive science and generative social science has yet to be leveraged for the purposes of improving our understanding of both.

The Resolution Thesis can be understood from multiple perspectives. From the generative social science perspective, the Resolution Thesis means that the representations of agents in social simulations should be informed closely by cognitive science and relevant neurophysiological considerations. This runs somewhat counter to the principled adherence to simplification of the internal processing of simulated agents found in this literature, one that, in fact, served to show that complex social dynamics can be driven by simple behavioral rules of

[1]Cognition considered as theoretical models of human perception, thought and action that includes, broadly, explanations of emotion, motivation, and affect in addition to more tradtional domains of cognitive psychology and cognitive science; one could arguably use the term *psychological first principles* as equivalent

agents. However, more recently, there are efforts in the generative social sciences that acknowledged that closer ties to the psychological and neurophysiological underpinnings of human behavior may yield benefit. Epstein's neurocognitive approach is a notable effort in this vein[REF, Epstein 2014]; there are other related approaches (e.g. REF 19-22 from BRiMS). These efforts notwithstanding, there remains a large gap between them and the implementation of models from cognitive science and psychology, not necessarily in principle, but in practice. It is worth noting that there are some threads in cognitive science that align with this perspective of the *Resolution Thesis*. In particular, Ron Sun's push for multi-agent systems based on cognitive agents and using cogitive first principles as the basis for the social sciences[5]; other work in this vein exists[1]. The relatively new field of computational social psychology is clearly relevant as well [6] as well as the computational organizational theory approach[4].

From the view of cognitive science, the thesis means that patterns of organization (e.g., information flow on the internet [REF; FROM SOCIALSIM], clustering of behaviors in a community[REF CHRISTAKIS]) at the social and organizational level should inform cognitive models when appropriate[2]. In other words, these patterns should be included as convergent evidence for a theory or model of cognition. At first consideration this notion may seem hard to fathom because the implications of cognition for the structure and dynamics of social systems are little understood from the cognitive science perspective[3]. Without an explanatory scheme that links facets of the cognitive model to aspects of social organization, how do we interpret the convergent evidence from social systems? The work mentioned above w.r.t. the cognitive first principles within the generative social science approach [5, 4, 6, 1] begins to put in place a better understanding of the implications of cognition on social systems, but it is nascent.

A third and more general view is that the Resolution Thesis is about human social systems. An understanding of any of these levels of scale is dependent, to some degree, on an understanding of the others. In effect, the notion of convergent evidence as originating, in part, from other levels of scale, applies to all levels of scale. The implication is that we should leverage information across scale in an iterative and synergistic way, if not simultaneously.

The Resolution Thesis, despite sounding both reasonable and practical at face value, is opposed by argument from both the cognitive sciences and generative social sciences. Simon's notion of nearly decomposable systems–that the temporal dynamics of adjacent levels of scale, in most systems, are little correlated–suggests that we can understand well the dynamics at each level of scale independently of the others(*REF, SIM 1962; see P. Anderson, 1972, for similiar argument in phyical systems). The KISS principle (keep it simple, stupid) used heavily in the generative social sciences, is clearly akin to Simon's

[2]The appropriateness may not be easily determined; for social cognition it may be obvious, but for perceptual domains, e.g., visual categorization, it is not obvious but also highly possible [**?**]

[3]Anderson's Relevance Thesis, a rare exception, reasons about how cognition may have implications for social organization.

notion, and is bolstered by its early wins in understanding the behavior of social systems using simple, non-cognitive agents (REF 15-17 from BRIMS). In cognitive science, Newell, in considering the time scale of human behavior, suggested that the social band ($> 10^4$ seconds, representing social systems and organizational behavior) is characterized to be weak in strength in the sense that it may not provide computations in a systematic way (relative to lower temporal bands, e.g., cognitive and neural processes).

These counter arguments notwithstanding, our working assumption is that the degree to which the *Resolution Thesis* is useful is an empirical issue. The state of the art in technology, computing and the their tight coupling to the current social mileau affords, we think, the testing of the Resolution Thesis. To this end, we've developed the *Reciprocal Constraints Paradigm* (henceforth *RCP*), a methodological approach that might lay a foundation for fruitful social simulation and an understanding of the implications for theory and models across levels of scale.

2 The Reciprocal Constraints Paradigm

The value of the *RCP* does not lie in precise formal presecription, but in providing a scaffold for building understanding of social systems, and, possibly providing a deep understanding of the interdependencies among levels of scale.

Figure 1 shows the four primary components of the *RCP*: a cognitive system with potential ties to neurophysiology, a social system, and the constraints between levels of scale. We assume that social systems and cognitive systems are derived from and exhibit an abstract set of first principles & properties, called S and C, respectively (e.g., theoretial entities, experimental paradigms, patterns in empirical data in respect to the discipines that address a particular level of scale). Defining S and C will depend on the social system or cognitive system of interest. The notion of constraint simply refers to the use of information from S & C to inform one another in a principled way. An upward-constraint refers to information from C informing S; downward-constraints reverse this relation.

A primary example in C is the set of allowable algorithms A such that $a \in C$. That is, A defines algorithms that are grounded in and thus recognized by work in cognitive science and psychology. Primary examples in S are the social structures, channels of information, and dynamic aggregate signals of behavior (e.g., a distribution of degree in a social network over time) that characterize a social system, much of which is formalized using graph theory/network science. It is important to emphasize that within S are notions regarding the behavior of agents, not only social structures.

A central assumption in the RCP is that cognitive systems and definitions of agent behaviors in social systems are meant to represent human information-processing capacities that can be described as mathematical functions. [**?**, **?**][4]. Thus, in C we can define a cognitive system as $\psi_{ct} : I_{ct} \rightarrow \psi_{ct}(i)$ where I_{ct} is the

[4]This is equivalent to Marr's computational level; we will use Marr's computational, algorithmic, and implementation levels of description[**?**, **?**] throughout this paper.

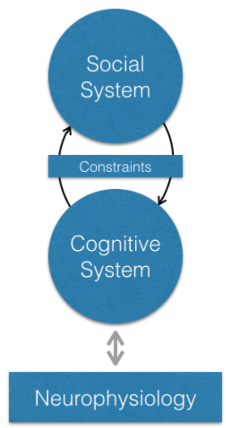

Figure 1: The four components of the *RCP* are a cognitive system with potential ties to neurophysiology, a social system, and the constraints on each one from the other. In the *RCP* social systems and cognitive systems are assumed to be derived from and exhibit an abstract set of first principles & properties, called S and C, respectively. Also captured here is the potential for integrating neurophysiolgical considerations into the cognitive system when appropriate; these may prove as essential for some social systems (the grey two-headed arrow indicates this potential). The notion of constraint refers to the use of information from S & C to inform one another in a principled way.

set of allowable inputs and $\psi_{ct}(i)$ is the output; in S we have a corresponding agent definition as $\phi_{at} : I_{at} \rightarrow \phi_{at}(i)$ where I_{at} is the set of allowable inputs and $\phi_{at}(i)$ is the output for an agent[5].

2.1 Applying the Reciprocal Constraints Paradigm

In practice there are multiple approaches available for application of the RCT, but what unites them is the study of a human social phenomena, either defined at one level of scale or at multiple levels of scale. Naturally, the first step is to identify a social phenonmena of interest, a task that is inherently tied to one's perspective. If the perspective is largely cognitive, then the focus would most likely be on understanding the psychological processes, representations, etc. in relation to social systems, but informed in some way by S. Another perspective, at the social scale, would dictate a concern with the social structures and dynamics of the social system (multiple humans interacting) with some degree of constraint from cognitive first principles. These two perspectives are both what we call single-scale approaches to the *RCP*. Of course, one could take a multi-scale perspective that draws from both and likely depends on a simulation approach that captures aspects of C and S in one runnable system.

We will address the obvious issues and difficulties in applying the *RCP* after providing a description of some potential methods for application. The goal in this section is simply to express what it might look like to apply the *RCP*.

*BUT KEEP IN MIND, RECIPROCAL CONSTRAINTS PARADIGM MEANS RECIPROCAL WHICH IS DIFFICULT AT ONE LEVEL OF SCALE.

2.1.1 Single-Scale Approaches

The single-scale approach of the *RCP* aims to elucidate or refine a model at one level of scale by using some information from another level. The essence of this approach is captured the following question: what are the implications of some information from another level of scale for the scale of interest.

Consider the cognitive scale. One potential approach, then, for implementing the *RCP* would be to map general properties of cognitive systems with properties of social systems for the purpose of identifying some of the implications of social structure and dynamics for cognition. For example, a central finding in cognitive science and psychology is that the products of learning mechanisms are sensitive to the order in which information is presented to the system (e.g., a human's generation of a mental representation of a set of stimuli depends on the order of exposure to the stimuli). The application of the RCT, in this case, means inferring how social features may systematically affect the temporal ordering of stimuli for human contexts. (This seems, almost axiomatically, dependent upon some properties of S if S contains graph G where $V(G)$ and $E(G)$ are the agents and information channels, respectively.) Off-hand, there seem to

[5] Social and cognitive systems may define parameters regarding variability among a set of agents; this is not reflected here.

be candidate examples from the literature regarding human socio-technical networks(e.g., [REF BEARMAN; REF BARABASI WEB]), but what this means precisely would depend on the social phenomena of interest (e.g., early language development may depend on a different $G \in S$ compared to racial stereotypes or large-scale population biases in attitudes).

At the social level of scale, an obvious approach is an analysis of the degree to which ϕ_{at} compares to any $\psi_{ct} \in C$ for the purpose of clarifying to what degree a social agent seems to be aligned with cognitive first principles. For the case in which ϕ_{at} and ψ_{ct} are formally well defined, this might be relatively straightforward[6]. But, there will certainly be cases for which this is not true[7].

*WHAT ABOUT THE THRESHOLD MODEL AS AN EXAMPLE? For example, imagine that the closest ψ_{ct} is only defined in terms of an experimental paradigm, theoretical apparatus and the interpetation of data resulting from application of the experimental paradigm[8]. One appraoch for comparison in this case would be to explore the behavior of ϕ_{at} by simulating an experimental paradigm that is isomphorphic to that used to define ψ_{ct} (assuming ϕ_{at} is defined algorithmically). In essence, this is like running an psychological experiment on artificial agents, an approach that is akin to simulation in the psychological sciences. The output of these experiments could be compared to the patterning and dynamics of human performance in C.

2.1.2 Multi-Scale Approach

The multi-scale approach is simple in principle: build a simulation platform that simultaneosly captures essentail aspects of both S and C (e.g., an agent-based model of cognitive agents). The upward-constraints would mean a type of substition of ψ_{ct} for ϕ_{at} (substitution might be one approach; for others this might be more akin to a focus or commitment to ψ_{ct}). The downward-constraints, generated by some measure of how well the simulated social system matches S as defined by the pheonomena of interest, would serve as a signal that would suggest modifications to S, C or both.

The apparent simplicity of the multi-scale approach is misleading. The details will include some non-negotiable decision points; free parameters will abound. The primary issue is, naturally, to decide where to begin? Does one start with S and import aspects of C or the reverse? One might also consider starting from scratch and bringing together properties from S and C in a new way. Let's explore these options via a concrete example which will, we hope, surface some of the key decision points.

Imagine we're interested in the patterning of obesity by race/ethnicity, a phenomena of interest with key social and policy aspects (e.g., cultural attractors, social and shared-enviromental influence, spatial patterning co-occuring

[6]Potential methods for such a comparison would, idealy, focus not only comparison of input/output functions but also the nature of the runnable algorithm

[7]Most theoretical entities in psychology are not formalized in precise mathematical or computational terms.

[8]Some might argue that under these conditions, the notion of a functional mapping ψ_{ct} is nonsensical

with racial segregation and residential mobility) and, in fact, a pre-existing set of social simulations from which to draw[REF review paper on simulation of obesity]. Assume we adopt one of the existing simulation approaches, none of which incorporate much in C. Given this case, one approach would be to substitute ψ_{ct} for ϕ_{at} while keeping the other aspects identical to the original simulation. To do this, however, because both ψ_{ct} for ϕ_{at} are input-output functions, there is a minimum requirement to find a reasonable substitution such that both I_{ct} and $\psi_{ct}(i)$ are similar to both I_{at} and $\phi_{at}(i)$. This is non-trivial in the sense that there is a limited but variable set of candidate ψ_{ct} to consider. The field of obesity in the health behavior tradition (argueably a co-opting of social psycholgy first principles) has several candidates, each reasonably considered as different (e.g., $\{\psi_{ct_i}, \psi_{ct_j}, \psi_{ct_k}, ...\}$). This analysis may be difficult or may require that not all aspects of ϕ_{at} can be captured by substitution. This last issue notwithstanding, as long as the inputs I_{ct} can be defined in the context of the properties of S (and even if ouputs don't match well), we could still have a runnable simulation.

Next, we must consider the downward-constraints in this scenario. Imagine that along with a pre-existing social simulation from which to co-opt, come empirical data, judged of adequate quality, that could be used to compute an objective function w.r.t. the simulations accuracy. This signal, then, would serve as the downward-constraint, albeit in a relatively ill-structed way taken alone. However, in this scenario, one could imagine systematic variations in the objective function w.r.t. properties of ψ_{ct} surfaced through sensitivity analysis/Monte-Carlo methods (or another method of search through the space defined by ψ_{ct}. This leads to another decision point, how do we judge the validity, if you will, of the space spanned in ψ_{ct}, given this scenario. Surely, some would rightly argue that not all parameters in ψ_{ct} should be free to vary on theoretical grounds[**?**]. Further, unless the properties in S captured in the simulation completely reproduce the empirical data (within a reasonable degree of tolerance), one needs to consider, given the objective function, whether to vary some components in the simulation that represent something in S. In other words, in the above scenario it is probably not good practice to only focus on changing C, so decsions must be made in this regard.

The above scenario is but one approach, one that is *fixing S and importing C*. But what if we *fix C and generate S* instead? What does this look like? We will stay with the obesity example from above, but change the scenario such that we don't know about any simulation approaches that represent mainly S. Instead, imagine that S is composed of a set of population-level empirical studies, some of which include information on social networks and the built environment, and some theoretcial statements about peer-influence on social networks. Furthermore, similiar to the scenario above, there is a limited set of candidate ψ_{ct} to consider ($\{\psi_{ct_i}, \psi_{ct_j}, \psi_{ct_k}, ...\}$) in the health behavior tradition and, likely, other aspects of C that may not be represented in them (e.g., categorization, learning, and attractor dynamics). Further, only one of these candidates[9], ψ_{ct_i},

[9] There do exist a small handful of computationally implemented health behavior theories,

is in a computational formalism; we decide that this will be our starting point. Analysis of ψ_{ct_i} reveals that it captures the learning and on-the-fly formation of attitudes towards specific health behaviors (considered a precursor to behavior); it is composed in a general manner such that it applies to virtually any health behavior, and; it is empirically grounded using traditional health behavior theory measurment techniques (in one particular behavioral domain). All of this is useful, but some key components are missing. In particular, the model is mute with respect to the generation of social strucuture and dynamics (e.g., decision making for initiating/disolving relationships; social influence mechanisms in terms of how others' behavior or attitudes can potentially serve as the input to the system). Thus, at minimum, some decision points arise about network structure and the mechanism of socialinfluence–e.g., one could implement a static network topology from human empirical data; in terms of social influence, we might assume that what is spreading are attitudes and that exposure to others' attitudes can serve as input to an agent and, further there is a knowable function between attitude and behaviors relating to obesity, e.g., energy balance behaviors relating to caloric intake and use). (As an aside, we've implemented models of this sort [**?**].) At this point, we could build out further social structure and dyamics in relation to the problem of interest, racial/ethnic distributions in obesity, and thus would require further decision points. However, at some point, we need to run the simulation and determine how to apply the downward-constraint. We explore this next.

Once the simulation is runnable, the downward-constraint could operate, as describe above, by computing an objective function w.r.t. the accuracy of the simulation compared to extant empirical data. Here, the issues are mainly the same as in the *fixing S and importing C case described above.*, with an important, subtle difference. In this case, although, depending on the quality and provenance of the data sources (to mean, roughly, which level of scale do they originate from), there will be decision points in terms of whether to modify C or S, the focus should, at least initially, be on C. In the limiting case, imagine that the only data we have is from C, e.g., some behavioral studies that reveal something about the learning and formation of attitudes about obesity related behaviors and the relation between attitudes and the behaviors. All of these studies, further, were conducted at the individual-level and reveal little about social structure and dynamics. In this case, the objective function would be computed with respect to how well ψ_{ct_i} captures the variation in these data alone. In this case, it is possible that aspect of both C and S could be modified in the simulation to optimize the objective function. What is interesting about this limiting case is that C may still depend on S. Getting the S right would be essential.

These two examples are largely hypothetical. The value, we hope, is to provide some sketches of what the *RCP* might look like in practice. Clearly, there are many issues that are raised, even by these sketches, let alone by the general notion of the *RCP* and the *Resolution Thesis*. We address some of the

see[**?**] for a review

obvious of these issues next.

2.2 Issues in Applying the Reciprocal Constraints Paradigm

Both single- and multi-scale approaches to the *RCP* are fraught with difficulty in application.

2. FREE PARAMETER PROBLEM AND Acountable modeling, mere parameter optimization?:

3. What if S is dependent on C?

4. Many competing cog models for any phenomena ... Issues of the right cogntive model, lots of cog models....

5. Downward constraints in multi-scale approach....what precisly is changed

6. quality of the data...some are quite veridical on some dimensions, but may be misleading on others...e.g., tweets are tmeporal and veridical in terms of content, but inference about meaning, motivation and intention may not be possible. data issues in general...the case when there is little data or there is a patchwork of data.

The second kind is to abstract some details from full-fledged cognitive models, but adhere to a principle of *Accountable Modeling***??**.

Have data at both levels or just one...re

3 An Example of RCT Implemented

We now offer an examle of a multi-scale approach to the *RCP*. It is a highly stylized example that focues much on C using the cognitive architecture ACT-R with some properties of a specific social domain, online social coding in GitHub. In particular, [describe simulation sketch here].

It will be useful to introduce ACT-R before continuing to the simulations. ACT-R is a specific approach to modeling human cognition that invokes the idea of a cognitive architecture, a computational instantiation of the structures, mechanisms and represenations of the mind that are coherent with respect to another notion, that of a unified theory of cogntion[2]. A cognitive model that is developed using a cognitive architecture has a potential a constrains operations (if you will it constraints ψ_{ct} to limits of the human mind w.r.t. a unified theory of cognition standpoint. Cognitive models based on architectures are structured algorithms of ψ_{ct}, and, thus, not strictly normative; individual differences are captured naturally in cognitive architectures (e.g., working memory; expertise via domain knowledge). The generation of quantitative predictions given hypothetical situational domains provides rigor in estimating the quality of model in terms of empirical data.

ACT-R is composed of several dedicated modules(e.g., procedural and declarative memory, working memory, percetipn and action) that operate in parallel but asychronously vai capacity-limited buffer interfaces. Each module is composed of independent mechanisms, to include symbolic information processing structures that are combined with equations for the representation f specific regularities and phenomena (e.g., reinforcement learning, power law of practice, power law of forgetting). Further, ACT-R adapts to environmental structure via a number of learning mechanisms. In short, it provides an account of human informaion processing capability and limitations (e.g., inherent biases) precicely because of the coupling of computational mechanisms and limits on capacity (e.g., working memory, attention).

Using a cognitive architecture, versus any cognitive model, characterizes well what agents are supposed to do in the generative social science approach–adaptive behaviors that are coupled to the social and physical environment. ACT-R, in particular, has modeled a broad range of human behaviors to include language, complex decision making, experimental psychology paradigms, and dynamic task environments (see the ACT-R web site[**?**]). Of particular relevance for this paper are some efforts using ACT-R as the basis for agents in multi-agent social simulations–e.g, iterated two-player games, bot adversarial and cooperative [**?**]

In fact

3.1 Importing Neurophysiology

Base this on Stocco's paragraph from BRiMs and use it in the context of the ACT-R sim we did above. and

Also, add the paragraph from BRiMS on how neuro to cog is many to one and cog to social is one to one.

4 Discussion

Why not just Deep Learning...?

4.1 Moving Forward

From the single level of scale approach, One idea is to catalog social phenomena captured by psychological models and do a mapping of C to S. [10].

Downward constraints on neurophysiology? From neuro to cog to social to cog to neuro? Is this where it is headed? Stocco has a paragraph in BRiMS on the cog to neuro, use it.

References

[1] S Bhattacharyya and S Ohlsson. Social creativity as a function of agent cognition and network properties: A computer model. *Social Networks*, 32:263–278, 2010.

[2] Allen Newell. *Unified theories of cognition.* Harvard University Press, 1990.

[10] In the limit, the social features and the cognitive features have co-evolved on an evolutionary time scale

[3] et al Orr, Mark G. Sbp-brims multi-scale resolution paper. *LNCS*, X:a–b, 2018.

[4] Michael Prietula, Kathleen Carley, and Les Gasser. *Simulating organizations: Computational models of institutions and groups*, volume 1. The MIT Press, 1998.

[5] Ron Sun. *Cognition and multi-agent interaction: From cognitive modeling to social simulation*. Cambridge University Press, 2006.

[6] Robin R Vallacher, Stephen J Read, and Andrzej Nowak. *Computational Social Psychology*. Routledge, 2017.