(Chosen article: Griffiths et al.)

In this paper, the authors argue that the principle of rationality can not only be applied to Marr’s computational level as previously believed, but can also be pushed down towards the algorithm level, serving as a bridge from the computational level to the algorithm level. There are two reasons for exploring the intermediate level: (1) The recent popularity of rational models of cognition and Bayesian analysis has resulted in computational level theories, whereas, traditionally, cognitive psychology focuses on algorithmic level theories; therefore, we need to explore how these two lines of research relate and connect to each other. (2) There have already been some unintended efforts exploring the intermediate space between computation and algorithm level in the history of psychology; these efforts often consider cognitive constraints when defining computational-level models, while the authors think that these constraints should be considered independent of the computational level, and further efforts are therefore needed to clarify and organize existing research. I agree with the authors that intentional efforts of bridging Marr’s different levels are in urgent need. Without studying computational level, answers to the “how” question that traditional cognitive psychology is seeking will not be very insightful since they are only providing a description of how our mind might be working, but do not provide information on why this is the case; similarly, without relating to algorithm level, traditional computational level accounts may sound like too perfect a “dream” to really make sense under numerous constraints of the real world. Connecting the two levels are important for the continuous development of cognitive science; it is also doable given all the existing efforts the authors have mentioned.

The new approach the paper proposes is that, instead of building constraints into computational-level theories, computational-level theories should be defined without considering limitations on its execution, and we can later explore the consequences of these limitations, which will bring us closer to an algorithm-level theory. This argument also makes sense since this is exactly how Marr’s levels are defined. The computational level should always aim for the optimal solution; constraints only exist when these solutions are being performed on some real-world medium at lower levels (algorithm and implementation levels). If a computational-level theory takes its own constraints into consideration, this theory would no longer be at the computational level since it is shaped by something not even exist at this level.

The authors argue that rationality can be a fundamental tool for building the bridge between computation and algorithm levels during this process if resource limits are taken into consideration. Two approaches are discussed in the paper: (1) Asymptotic rationality in rational process models; (2) Rational use of finite resources (e.g., time limits) in resource-rational models.

The rational process models are process models generated from a computational-level solution to a problem of cognition using approximation algorithms like Markov chain Monte Carlo and particle filters. These process models are rational because they are derived from computational-level accounts; they are also down to the algorithm level because the specific process models generated by approximation algorithms each have their own constraints in time, memory, and the conditions under which they succeed and fail. However, the shortcoming of this approach is that it is only a weak form of rationality, i.e., rationality by derivation, since the process models can be regarded as being “rational” only because they are derived from a computational-level rational account of cognitive processes, but there is no way to figure out which one of these models are optimal. I agree that this is a promising way of bridging the two levels – it offers a reliable way of converting computational-level theories to specific algorithms, different from previous research where the two levels are studied almost in parallel in different types of studies.

Therefore, the authors propose a new approach (resource-rational models) through which optimal models can be identified when a specific computational-level solution is used. The first two steps of resource-rational models are similar to rational process models: finding an optimal computational level solution and generating a family of algorithms that approximate this optimal solution. But resource-rational models take a step further: after a family of algorithms is generated, the algorithm with an optimal trade-off between approximation accuracy and resources like time is selected, and compared to human behavior for refinement. In this way, this selected model has taken both the computational and algorithmic levels into consideration, and it is also “rational” in its strong sense since it is optimal in accuracy-resource trade-off. This approach may look plausible at first glance, but it could be in danger of infinite regress. This approach presumes that an optimal trade-off between approximation accuracy and resources is what is happening to the human cognitive system at the algorithm level; but this presumption itself is actually a computational level claim, and therefore itself also needs to be grounded at the algorithm level before it can be regarded as plausible. But how exactly can this computational principle be grounded at the algorithm level, i.e., how can we determine which one among all the possible algorithms is the right one? – It is obvious that we have again gone back to the exact question that this principle is claiming to solve. Do we need another “meta-” optimal-trade-off-principle in order to solve this problem? Then how can this “meta-” principle be grounded at the algorithm level before we can be sure that this really is the case with human cognition? … This circular line of questions can go on and on and on, and we thus go into an infinite regress.