Computational psychiatry

https://doi.org/10.1038/s44159-024-00360-7

Computational approaches to mental illnesses

Theoretical understanding of psychopathology has always been reliant on phenomenology (that is, clinical observation). Although alluring given its simplicity, this descriptive approach has proven severely limited in addressing the diagnostic and ontological complexity of mental illness. A different approach based on computational science is gradually emerging as a radical alternative.

This computational perspective, as eloquently outlined by Quentin Huys and colleagues, draws a distinction between discriminative and generative modelling. The former includes statistical models (such as latent variable, network and machine learning models) that are already part of psychologists' methodological arsenal. These data-driven models are immensely useful in terms of revealing phenomena such as the reinforcing relations among psychopathology symptoms or the robust association between unemployment and addiction. However, these models fall short in terms of explaining those phenomena: because their data-generating mechanisms are somewhat general, they cannot speak to the precise processes that are (most) likely to have generated their fitted responses.

Generative (or computational) models have the opposite properties. In particular, generative models are uniquely crafted to instantiate specific data-generating processes that are theoretically meaningful. Accordingly, such models have greater explanatory power because they enable researchers to test (usually experimentally) precise hypotheses about the origins of particular psychopathologies. At the same time, these models are currently weak predictively, because their experimental paradigms have either low validity (they do not exactly evoke a specific computational mechanism) or low reliability (they do not yield sufficient between-subjects variability).

To outline advances in these and other computational matters, Huys and colleagues discuss how various formal models have been applied to psychological experiments to elucidate the processes that underlie mental illnesses. First, reinforcement learning models have been used to explain how emotional problems arise from mismatches between what people 'expect' (or desire) and what they 'get' in life. Second, probabilistic models have been used to formalize beliefs (via simple statistics such as the expectation and variation of probability distributions) and their maladaptive variants (for instance, polarized beliefs as skewed distributions). Third, dynamical systems have been used to study the network dynamics of psychopathologies at both the psychological (for example, existential tipping points) and neurobiological levels (for example, attractor states toward which neuronal populations gravitate). Fourth, game theoretic approaches have been used to understand the interpersonal problems in personality disorders, casting them as suboptimal ways of cooperating.

Finally, beyond elucidating pathogenic mechanisms, computational tools are increasingly being used to formalize the salutogenic mechanisms of psychotherapy. Specifically, psychotherapy is increasingly understood to operate through belief-updating mechanisms that either relax the precision (conviction) of maladaptive beliefs (through top-down cognitive change) or generate evidence for more adaptive beliefs (through bottom-up behavioural change).

> **"Unlike statistical models, which are all 'wrong', computational models are expected (or hoped, rather) to be 'right.'"**

The renowned status of computational science has prompted widespread enthusiasm for computational models of mental illnesses. Indeed, unlike statistical models, which are all 'wrong', computational models are expected (or hoped, rather) to be 'right'. Whether or not that is the case, though, remains to be seen.

Orestis Zavlis
Department of Psychology and Language Sciences, University College London, London, UK.
e-mail: orestis.research@gmail.com

Competing interests
The author declares no competing interests.

Original article: Huys, Q. J., Browning, M., Paulus, M. P. & Frank, M. J. Advances in the computational understanding of mental illness. *Neuropsychopharmacology* **46**, 3–19 (2021)

Related articles: Karvelis, P., Paulus, M. P. & Diaconescu, A. O. Individual differences in computational psychiatry: a review of current challenges. *Neurosci. Biobehav. Rev.* **148**, 105137 (2023); Smith, R., Moutoussis, M. & Bilek, E. Simulating the computational mechanisms of cognitive and behavioral psychotherapeutic interventions: insights from active inference. *Sci. Rep.* **11**, 10128 (2021)