Computational psychiatry

“(...) seeks to characterize mental disorders in terms of aberrant computations at multiple scales.”
Addiction as a Computational Process Gone Awry

A. David Redish

Addictive drugs have been hypothesized to access the same neuropharmacological mechanisms as natural learning systems. These natural learning systems can be modeled through temporally difference reinforcement learning (TDRL), which requires a reward-reversal signal that has been hypothesized to be carried by dopamine. TDRL learns to predict reward by dividing that neurotransmitter signal to zero. By adding a noncompulsively drug-induced dopamine increase to a TDML model, a computational model of addiction is constructed that over-acti- onates the reward system leading to drug craving. The model provides an explanation for important aspects of the addiction literature and provides a theoretical view point with which to address other aspects.

From reinforcement learning models to psychiatric and neurological disorders

Ting V Maia1,2 & Michael J. Frank1,4

Over the last decade and a half, reinforcement learning models have fostered an increasingly sophisticated understanding of the functions of dopamine and cortico-basal ganglia-thalamo-cortical (CBGTC) circuits. More recently, these models, and the insights that they afford, have started to be used to understand important aspects of several psychiatric and neurological disorders that involve disturbances of the dopaminergic system and CBGTC circuits. We review this approach and its existing and potential applications to Parkinson’s disease, Tourette syndrome, addiction, obsessive-compulsive disorder, and post-traumatic stress disorder. As these clinical and preclinical animal models are used to create new antipsychotic drugs, the approaches for the identification and prediction of new powerhouses will be for the continued growth of computational psychiatry and computational neuroscience.
Can ACT-R contribute the missing piece?
Intrusive Memories

> Emotionally charged memories that intrude in everyday life and prevent normal functioning
> Staple of several disorders (depression, OCD, ...)
> Particularly studied in PTSD
The problem: Not all patients are equal

Galatzer-Levy, Huangb, & Bonanno, 2018, Clinical Psych Review
ACT-R has an excellent model of memory!

\[
\frac{p(i|Q)}{p(-i|Q)} = \frac{p(i)}{p(-i)} \times \prod_{q \in Q} \frac{p(q|i)}{p(q)}
\]

\[
\log \left( \frac{p(i|Q)}{p(-i|Q)} \right) = \log \left( \frac{p(i)}{p(-i)} \right) + \log \left( \prod_{q \in Q} \frac{p(q|i)}{p(q)} \right)
\]

Total Activation \( A_i \) = Base-Level Activation \( B_i \) + Spreading Activation \( S_i \)

\[
A_i = B_i + \sum_i \log(t_i^{-d}) + \sum_i WS_{ji}
\]
Memory is rational and elastic
Incorporating emotions and memory

> Activation tracks need probability, $A_i \sim P(i \mid Q)$
  – Assumes all memories are created equal.
Not all memories are equal!

(One-shot learning at its finest)
Incorporating emotions and memory

- Activation tracks **need probability**, \( A_i \approx P(i \mid Q) \)
  - Assumes all memories are **created equal**.
  - Some memories are **intrinsically more important**!
- \( A_i \) should reflect a memory’s **expected value**, i.e.
  \[
  A_i \approx P(i \mid Q) \times V(i)
  \]
- Old idea (e.g., West, Larue, and my 2004-self)
In addition to slots, chunks are given a scalar quantity, $V$
- Insight: Emotions have “survival value” (Panksepp)
$V$ is actually normalized, given the rest of memories, so that $M(V) = 1$
Code at github.com/UWCCDL/PTSD
Implementation

\[
\frac{p(i|Q)}{p(-i|Q)} = \frac{p(i)}{p(-i)} \times \prod_{q \in Q} \frac{p(q|i)}{p(q)} \times V(i)
\]

\[
\log \left( \frac{p(i|Q)}{p(-i|Q)} \right) = \log \left( \frac{p(i)}{p(-i)} \right) + \log \left( \prod_{q \in Q} \frac{p(q|i)}{p(q)} \right) + \log (V(i))
\]

Total Activation \[ A_i \] = Base-Level Activation \[ B_i \] + Spreading Activation \[ S_i \] + Survival Value \[ V_i \]

\[
\begin{align*}
A_i &= B_i + S_i + V_i \\
B_i &= \sum_i t_i^d \\
S_i &= \sum_i WS_{ji} \\
V_i &= \log(V_i)
\end{align*}
\]
This model is still **rational**, but memories are **inelastic**
Abstract agent that simulates a few months of life
Accrues random memories at fixed intervals
Productions implement a **perceive/retrieve/act** loop
  – “Retrieve” replaces “Decide”
  – Not too different from Christian’s IBL models
Testing the model

> At predefined time $T_{PTE}$, a traumatic event $PTE$ is introduced
  – Random memory, but $V > 1$
> We can now track down how often PTE is retrieved out of context
Effects of PTE Value

Traumatic memories as a function of value $V$

Non-linear effect of trauma (chronic vs. resilience)
Disruption of cognitive functions

> Function: relevance of a retrieved memory $R$ to the current situation $Q$
> Measured as $\cos(R, Q)$
> Traumatic memories more disruptive
Biological interpretation

> The PTSD field is looking for **biomarkers**
> Two targets:
  – Amygdala/hippocampus circuit
  – Prefrontal/hippocampus circuit
Biological interpretation and findings

- **Greater amygdala** = greater effect of trauma (greater value $V_i$)
- **Greater hippocampus** = smaller effect of trauma (smaller decay $d$)
- **Greater prefrontal volume** = smaller effect of trauma (greater spreading activation $W$)
- **Greater working memory** = smaller effect of trauma (again, greater $W$)
What’s next

> More simulations are running!
> Predicting PTSD from brain activity (should provide $V$ and $d$ through neurometrics)
  – ACT-R as the missing model between biomarkers and outcomes
  – With Katie McLaughlin
> Application: Optimizing presentation of traumatic images to human raters
Thank you!

Code at: http://github.com/UWCCDL/PTSD