Cognitive Modeling of the Behavioral Effectiveness of Non-Pharmaceutical Interventions

Christian Lebiere, Carnegie Mellon University, (cl@cmu.edu) in collaboration with:

Kathleen Carley, Michael Martin & Konstantinos Mitsopoulos, Carnegie Mellon

Peter Pirolli, Bonnie Dorr, Tomek Strzalkowski, Adam Dalton,
Brodie Mather & Archna Bathia, Institute for Human and Machine Cognition

Mark Orr & Stephen Eubank, University of Virginia Biocomplexity Institute
From Information To Behavior

• Build on:
  • Theories of individual health psychology
  • Predictive computational cognitive models of behavior change

• Psychologically Valid Agent-based epidemiological models
  • More accurately predict the dynamics of behavior change
  • Response to NPIs government messaging, mass media, social media, information framing, etc
Cognitive Architectures and Human Behavior

- ACT-R provides constrained, principled framework to model complex human behavior
- Modules for knowledge, action selection, working memory, perception and motor actions
- Integrates symbolic knowledge and statistical adaptivity
- Reflects individual differences and emergent cognitive biases
- Accounts for training effects by modeling learning processes
Instance-Based Learning (IBL)

- Instance-Based Learning (IBL) models make decisions by generalizing previous situations.
- Leverage cognitive mechanisms for memory (decay, rehearsal, priming) and pattern matching (partial matching, blending).
- Individual biases reflect distinct experience history (knowledge).
- Additional biases from cognitive mechanisms (e.g., recency bias).
- Predict training effectiveness by diversity, size of instance base.

\[
A_i = \ln \sum_{j=1}^{n} t_j^d + MP \times \sum_{k} Sim(v_k, c_k) + \varepsilon_i
\]

\[
P_i = \frac{e^{A_i/s}}{\sum_j e^{A_j/s}} \quad V = \arg\min \sum_i P_i \times (1 - Sim(V, V_i))^2
\]

<table>
<thead>
<tr>
<th>Reward</th>
<th>Penalty</th>
<th>Probability</th>
<th>Action</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>-9</td>
<td>0.36</td>
<td>Attack</td>
<td>8</td>
</tr>
</tbody>
</table>

Event Representation (Context)

8 partial match
9 partial match

8 partial match
9 partial match

4.1 blending
Theory of Planned Behavior

- predict and explain wide range of health behaviors and intentions
- Social norms reflect normative beliefs toward desired behavior
- Attitudes reflect intentions toward execution of behavior
- Control beliefs reflect perception of ability to control behavior
Social Norms

- Attitudes reflect input from agent’s global network
  - Social/political affiliation and social network polarization
  - Transfer from other norms e.g. drunk driving
- Attitudes are modulated by local context
  - Prevalence of behavior in local environment
  - Immediate peer pressure from direct contacts
- Formalized as chunks encoding competing actions
  - Wear-mask vs not-wear-mask actions
  - Competing actions rather than action-valence better fits IBL approach
  - Competing base-level activation of alternative action chunks
Evolving Social (Media) Norms in CA and PA
Mask Wearing Predictions in 4 Large States

- Memory strength of norms abstracted from ratio of pro- and anti-mask Twitter hashtags

- Reflects power laws of recency and frequency

- 1 parameter estimated: blending temperature
  - suggests stronger social cohesion in NY and CA

- Validated against CovidStates data
Mask Wearing Aggregate Validation

- Aggregate Twitter data and unweighted mask wearing probabilities
- Temperature parameter estimated near neutral value (1.2)
- Better fit than single states
  - Limitations of noisy data
  - National nature of twitter
  - Unrepresentative sample
  - Other behavior factors

![Twitter-Based Predictions vs Actual](chart.png)
Control: Damped Oscillations of Infection Rate

• Damped oscillations in Rt observed at all levels
• Attitude Representation
  • Context-free norm chunks
    • Wear vs don’t wear
  • Contextual (Rt) reactions
    • “Fear” vs “hope”
  • Chunk activation strength
• Context estimated from temporal aggregation

• Match context -> action
  • Norm damping effect
  • Contextual oscillations
Influence of Attitude on Rt

- Strengths of pro- and anti-mask attitudes represented as distinct activation of corresponding beliefs
- Stronger impact on asymptote of infection rate than initial response
  - Due to effect of natural reactions ("hope" and "fear") that are triggered regardless of attitude by the current state of the pandemic
- Population as heterogeneous mix of attitudes activated at different points, moving between curves
Dynamics of Attitudes and Infections

- Feedback delays in dynamical system control induce oscillations
  1. High mask wearing compliance reduce initial infection rate
  2. Lower infection rate leads to relaxing in mask wearing
  3. Lower mask wearing probability leads to rising infection rate
  4. Rising infection rates leads to increase in mask wearing
- Amplitude of oscillations decrease over time but can rise with external shocks (mandates, weather, events)
Behavioral Attitudes: Increasing Polarization

- Mobility reduction decreases over time
- Differences tied to socio-political orientation amplifies over time
- Fundamental task is decision making under uncertainty (safe vs risky)
- Well-studied problem can lead to dynamic effects such as cognitive biases

<table>
<thead>
<tr>
<th>Problem</th>
<th>Risky</th>
<th>Safe</th>
<th>Proportion of risky choices (R-rate)</th>
<th>Average number of samples per problem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.3</td>
<td>0.96</td>
<td>-2.1</td>
<td>0.20</td>
</tr>
<tr>
<td>2</td>
<td>-0.9</td>
<td>0.95</td>
<td>-4.2</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>-6.3</td>
<td>0.30</td>
<td>-15.2</td>
<td>0.60</td>
</tr>
<tr>
<td>4</td>
<td>-10.0</td>
<td>0.20</td>
<td>-29.2</td>
<td>0.85</td>
</tr>
<tr>
<td>5</td>
<td>-1.7</td>
<td>0.90</td>
<td>-3.9</td>
<td>0.30</td>
</tr>
<tr>
<td>6</td>
<td>-6.3</td>
<td>0.99</td>
<td>-15.7</td>
<td>0.35</td>
</tr>
</tbody>
</table>

I. Erev et al. A Choice Prediction Competition
Impact of Messaging on Behavior Adoption

- Previous behavior presented as risky
- Alternative behavior slightly less desirable
- Messaging competes with experience
- No confirmation bias storage of expectation
- Standard parameters
Integrating 3 Levels of Modeling

• Pro- and anti-mask wearing norms
  • Driven by prevalence of social norms but context-free

• Attitude-driven reactions
  • Context is estimate of pandemic expansion or contraction (~Rt)
  • Driven by affective reactions such as hope and fear

• Decision-making under risk and uncertainty
  • Driven by combination of risk messaging and direct experience
  • Reflects sampling dynamics, e.g., risk aversion and confirmation bias

• Combine reactive, affective and reflective decision making
  • Generate expectation -> capture habituation effects
  • Generate action to meet context and expectation
  • Activation strength of each structure modulates its effect
    • Reflect individual dispositions and exposure to information environment
Integrating Multiple Factors

• Modeling previously unseen pattern:
  • starting from a stable plateau (Rt~1) of February-March
  • gradual (quasi-linear, not exponential) increase in late March
  • plateau about 40-50% above original level in early April
  • decrease symmetric to original increase in late April through May

• Pattern reproduced at national, state and local levels
  • Largely independent of local characteristics

• Pattern present in cases, hospitalizations and deaths with usual time lag
Cognitive and Biological Mechanisms

• Using four concurrent biological and cognitive mechanisms

• Two external (pharmaceutical/biological) mechanisms
  • gradual introduction of vaccines that proportionally reduces Rt from a roughly linear decrease in the Susceptible population
  • sigmoid increase of factor ~0.4 in R0 (and consequently Rt) resulting from the spread of the more contagious British variant

• Two internal (cognitive/behavioral) mechanisms:
  • fairly sudden relaxation of behavior triggered by the introduction of vaccines and seasonal/official relaxation
    • Expressed in first level of cognitive model, i.e. norms, in terms of proportional increase of $A_d$ activation level
  • gradual tightening in behavior resulting from the original increase in cases
    • Expressed at the second level of the cognitive model, i.e., higher $A_f$ activation resulting from closer match to increasing Rt
Phases of Model Behavior

Near-exponential explosion of British variant and behavior relaxation overcomes gradual vaccination

Behavior tightening and increasing vaccinations stabilize against British variant

Early, slow increase in vaccinations and British variant, stabilized by cognitive dynamics

Vaccinations approaching herd immunity and behavior tightening overcome British variant saturation
Future Work

- Integrate 3 model levels & time scales of dynamics
- Improve fidelity of belief stance representations
- Integrate effect of diverse stances in model dynamics
- Integrate external events in control dynamics
- Extend model to account for vaccination decisions
- Integrate behavior model into large-scale ABM
Acknowledgements

This Work is supported by the National Science Foundation Grant No. 2033390 and the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via 2020-20092500001.