SAND2011-4593C

Applications of Self-Correcting Hybrid Causal-Learning Systems to Opinion Dynamics on Networks

Alexander Outkin and Robert Glass

June 27

ICCS 2011, Boston, MA

Motivation

- Social systems can be represented as:
  - Engineered component - *interaction structure and rules.* Enforces constraints and processes.
  - Decisional component - *agent's strategies.* Drives the system.

Networks and Markets

- Agent-based models can be thought of as representing the causal structure and mechanics of the world
  - Social networks: dynamics is driven by agents decisions, within social or technological constraints
  - Markets: agent's have heterogeneous strategies, operate within rules, regulations, and existing infrastructure

Question

- How predictive power generated by understanding of causal mechanisms can be incorporated into real-time decision-making?
- Difficulties:
  - Social systems are only partially observable - thus matching output of causal model to observed data is difficult
  - Both strategies and causal structure of the system can change abruptly

Approach

- Create causal (agent-based) model(s) of the system
- Enable incorporation of information from multiple models into a learning framework and evaluation of individual model contributions
- Enable causal models partial re-calibration as new data arrives

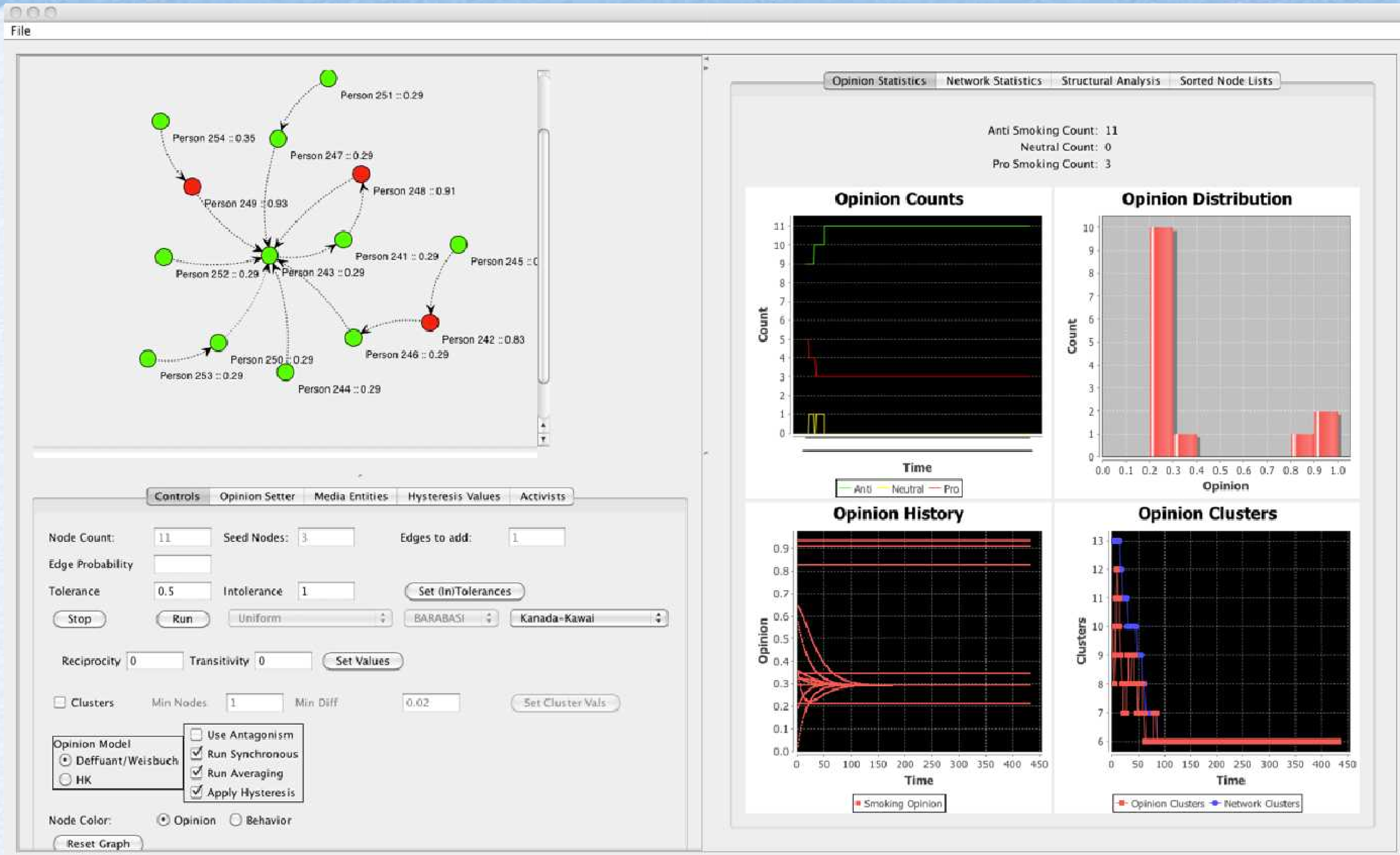
Opinion dynamics

- Real-world applications: identify a (set of) causal models that predict the future system performance and response to interventions
- Test case: create a randomly generated opinion model and underlying network and apply the learning framework to predict its dynamics

Test model

- Developed by Sandia CASoS initiative
- Network of agents who share and affect each other opinions
  - Weisbuch updating rule for example
- Network topology may vary: Barabasi, random, ring...

Test Model



File
Person 251 :: 0.29
Person 254 :: 0.35
Person 247 :: 0.29
Person 248 :: 0.91
Person 249 :: 0.93
Person 241 :: 0.29
Person 245 :: 0
Person 252 :: 0.29
Person 243 :: 0.29
Person 242 :: 0.83
Person 246 :: 0.29
Person 250 :: 0.29
Person 253 :: 0.29
Person 244 :: 0.29
Controls
Opinion Setter
Media Entities
Hysteresis Values
Activists
Node Count: 11
Seed Nodes: 3
Edges to add: 1
Edge Probability
Tolerance 0.5
Intolerance 1
Set (In)Tolerances
Stop
Run
Uniform
BARABASI
Kanada-Kawai
Reciprocity 0
Transitivity 0
Set Values
Clusters
Min Nodes 1
Min Diff 0.02
Set Cluster Vals
Opinion Model
Deffuant/Weisbuch
HK
Use Antagonism
Run Synchronous
Run Averaging
Apply Hysteresis
Node Color:
Opinion
Behavior
Reset Graph
Opinion Statistics
Network Statistics
Structural Analysis
Sorted Node Lists
Anti Smoking Count: 11
Neutral Count: 0
Pro Smoking Count: 3
Opinion Counts
Count
Time
Anti
Neutral
Pro
Opinion Distribution
Count
Opinion
Opinion History
Opinion
Time
Smoking Opinion
Opinion Clusters
Clusters
Time
Opinion Clusters
Network Clusters

Experiment Parameters

- What we can observe:
  - Average behaviors
  - Number of nodes (more or less)
  - Some topological or connectivity parameters
- What we cannot observe:
  - Individual strategies
  - Exact topology
  - Opinions or at times behaviors for individual agents

Learning Framework

- A set of inputs $X = x_1, \dots, x_K$. Those can represent past historic observations of the system's behavior.
- A set of outputs $Y = y_1, \dots, y_L$. Those represent system responses to the inputs.
- A set of interaction and correlation models: $M = m_1, \dots, m_N$, each representing a (sub) set of the systems behavior.
    - Each model m_i has a set of inputs x_{ij}, where $i \in K_i \subseteq \{1, \dots, K\}$.
    - Each model m_i has a set of outputs: y_{il}, where $l \in L_i \subseteq \{1, \dots, L\}$.
    - A goal of individual models is to represent its outputs as a function of inputs.
    - Additionally, a goal of interaction model is to represent the causal structures and interactions that give rise to its dynamics.
    - We calibrate the individual interaction models to their individual inputs and outputs.

Learning Framework

- A learning model, which is a function F of outputs of all models, whose goal is to predict the model outputs.
  $Y(t+1) = F\big(m_1(t), \ldots, m_N(t)\big) + \epsilon(t)$
  Compare this with a standard linear model, where
  $$Y(t+1) = \sum_i w_i x_i(t) + \varepsilon(t)$$

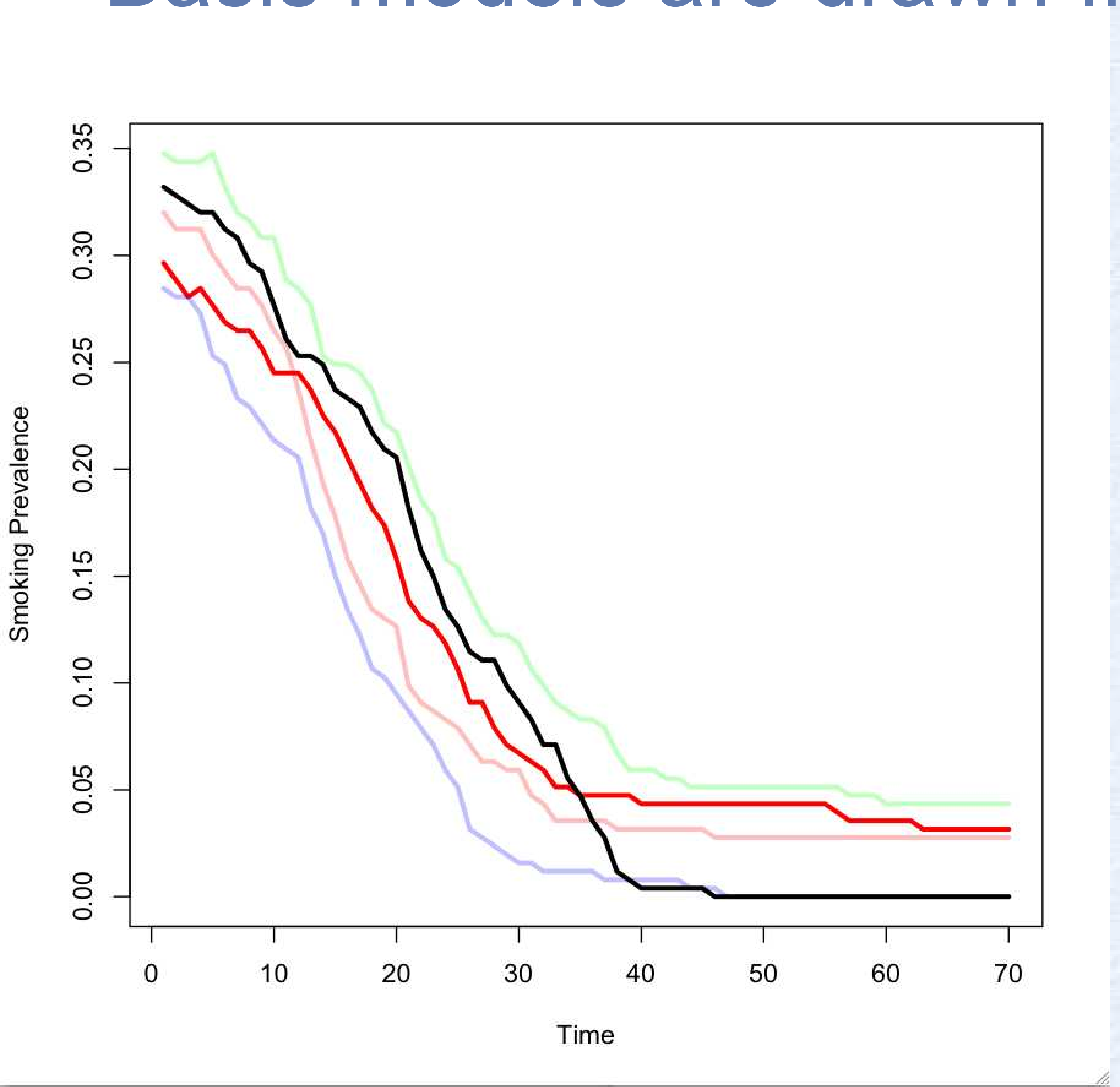
- In an online setting a key question is also how to re-calibrate the causal models with the newly arrived data

Experimental Setup

- Create a single instance of a network
  - Some of its parameters and initial values are randomly generated. Their actual values are not used for creation of “basis” models
- Create a set of families of models using known parameters:
  - Average opinion
  - Number of nodes
  - Information on network structure

Experimental Setup

- Example objective function – red
- Basis models are drawn from "same" distribution



Off-line Estimation

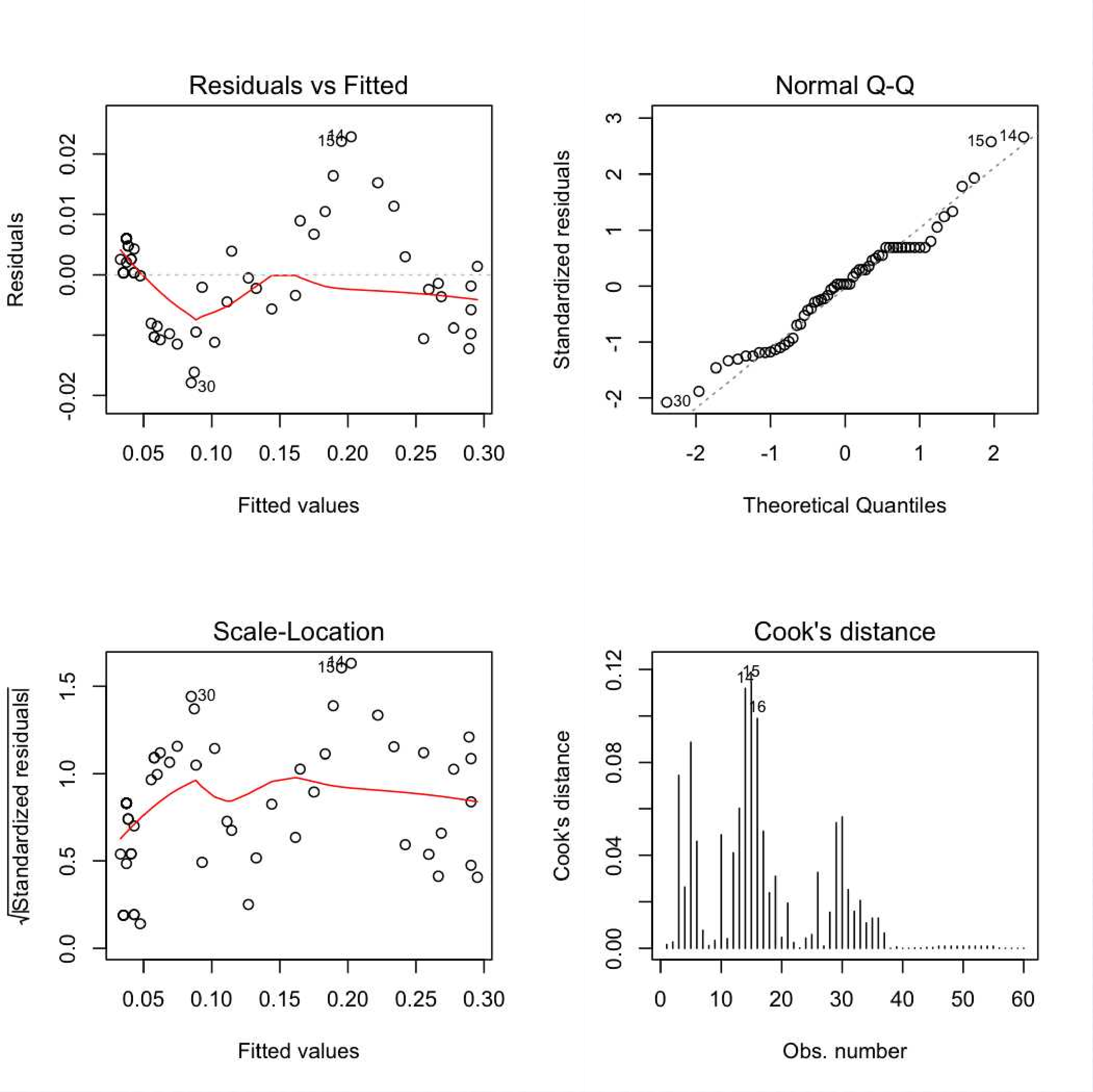
- Two randomly generated models are sufficient to represent the test model output when the "learning" function is a simple linear regression

$$Y(t+1) = \sum_i w_i m_i(t) + \varepsilon(t)$$

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
Experiment.1   0.31113    0.03716   8.373 1.46e-11 ***
Experiment.2   0.56198    0.02927  19.197  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.008718 on 58 degrees of freedom
Multiple R-squared: 0.9968, Adjusted R-squared: 0.9966
F-statistic:  8914 on 2 and 58 DF,  p-value: < 2.2e-16
```

Diagnostics

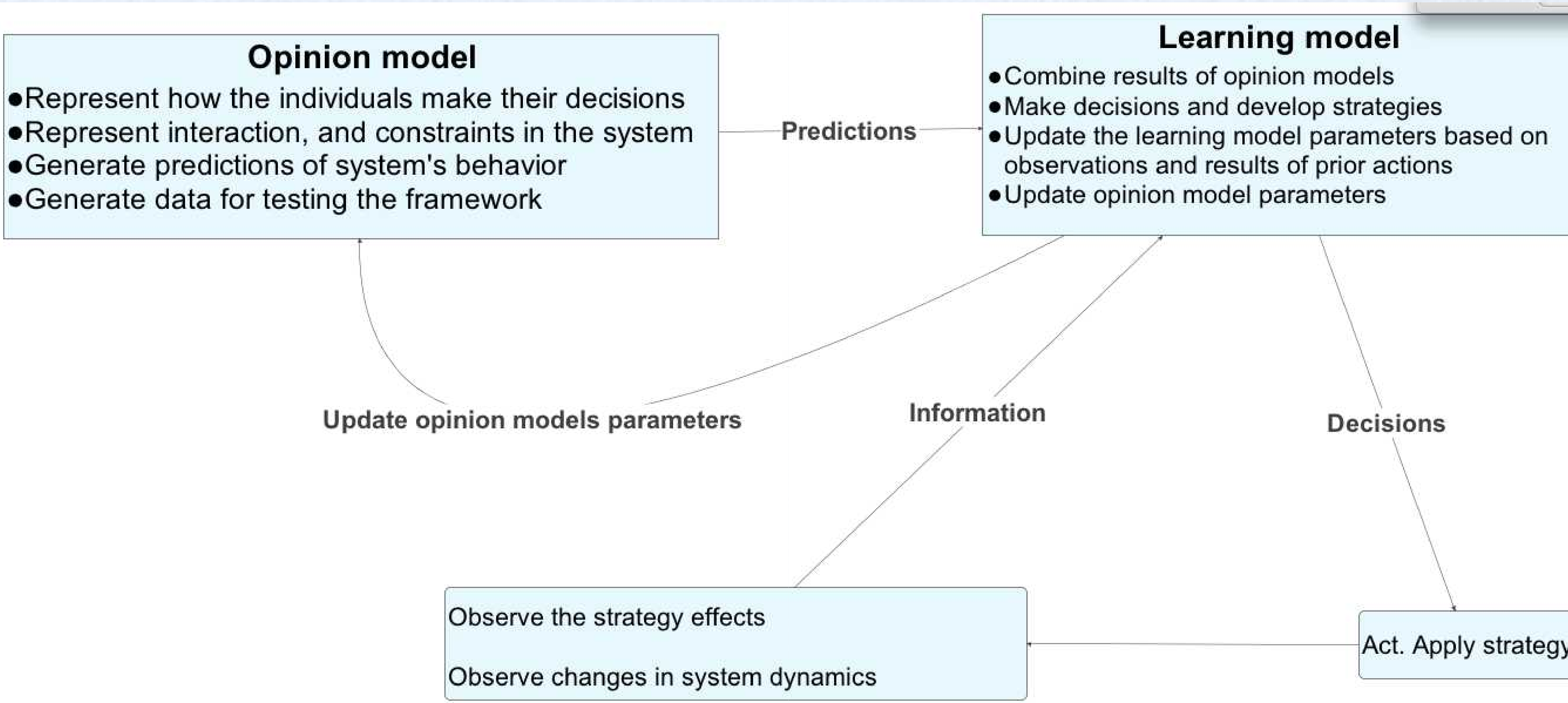


On-line Notes

- In on-line or real-time setting, the key usefulness of the causal models is in the ability to select across the models based on newly arrived information and to re-calibrate the models
- The main tasks are:
  - Which of the “basis” models are closes to the observed data?
  - How identify when the basis model ensemble is insufficient: regime changes?

Conceptual Framework

- Act/Observe/Select/Recalibrate



Conclusions

- Connecting causal models to an on-line or off-line learning framework provides an ability to use incomplete causal information for prediction and ability to select across many possible causal models.
- Preliminary results in the opinion dynamics demonstrate ability to replicate the results of a causal model with randomly generated causal models
- Future directions include online estimation and model selection, and regime changes identification