Models and Experiments

Joshua Becker
Kellogg School of Management
Northwestern Institute on Complex Systems

Summer Institute in Computational Social Science – Chicago

June 21, 2019

MORE RESOURCES:

https://www.joshua-becker.com/resources

Case Study 1: The Wisdom of Crowds
Motivation & Background

- Forming accurate beliefs is critical to decision-making
- Common theoretical expectation: social influence undermines the wisdom of crowds
- Intuitive theory: diversity prediction theorem

Case Study 1: The Wisdom of Crowds *Theory-Data-Theory*

- Model: decentralized networks converge on mean belief
- Experimental data: decentralized networks become more accurate
- Observation: accurate people revise less
- Model: decentralized networks become more accurate

Becker, J., Brackbill, D., & Centola, D. (2017). **Network dynamics of social influence in the wisdom of crowds.** *Proceedings of the national academy of sciences*, *114*(26), E5070-E5076

Tipping Points in Social Conventions
Motivation & Background

- Equilibrium theories state that once established, conventions are stable

- And yet ample evidence exists of sudden social change

- Folk theory: critical mass

Tipping Points in Social Conventions *Theory-Data-Theory*

- Model: 10% group is sufficient to overturn established convention
- Experimental data: 10% is not sufficient
- Observation: people are strategic
- Revised Model: 25% is sufficient to overturn established convention

Centola, D., Becker, J., Brackbill, D., & Baronchelli, A. (2018). **Experimental evidence for tipping points in social convention.** Science, 360(6393), 1116-1119.

Why Experiments? (The Scientific Trifecta)

Observation

Model

Experiment

The Computational Social Science Trifecta

Observation
(Big Data)

Agent-Based
Model

Web
Experiment

The Limits of Causal Inference from Observation: Diffusion of Medical Innovation

ORIGINAL STUDY: Coleman, J., Katz, E., & Menzel, H. (1957). The diffusion of an innovation among physicians. Sociometry, 20(4), 253-270.

REANALYSIS ONE: Burt, R. S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. American journal of Sociology, 92(6), 1287-1335.

REANALYSIS TWO: Van den Bulte, C., & Lilien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. American Journal of Sociology, 106(5), 1409-1435.

What is causality? (Logical approach)

$A \rightarrow B$ "If A, then B" $\equiv$ sufficiency

$\neg A \rightarrow \neg B$ "If not A, then not B" $\equiv$ necessity

** counterfactual is key **

How do we demonstrate causality?

- B follows A *(sufficiency)*
- Counterfactuals *(necessity)*
- Randomization
- Temporality
- ……?

What is an experiment?

- **From the lab:** ceteris paribus study where one (or more) parameters differs between conditions

- **From probability:** any replicable procedure with well defined outcomes

Crucial Tests

- Theory 1: accurate people respond less to social influence for all beliefs

- Theory 2: extremists respond less to social influence for political beliefs

- Crucial test: in homogeneous networks…
  (1) increases error of average
  (2) decreases error of average

Why experiments?

(this slide intentionally left blank)

Why formal models?

What is a model?

- Statistical model
- Generative model
- Agent-based model
- Coupled differentials

What do (generative) models do?

- Explanation: how is it possible that conventions form?

Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, 57-84.

- Prediction: how does network structure impact innovation?

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, *52*(4), 667-694.

Models do not have to be "correct"

- **James Clerk Maxwell:** model of vortices of frictionless fluids… used to derive electromagnetic equations

"All models are wrong. Some models are useful."

—George Box

What makes a model agent-based?

- Simulation != agent-based
- Equations != not agent-based
- Agent-based is generative

How to design an agent-based model

- Step 1: model of individual behavior of interest
- Step 2: identify structural/environmental features of interest
- Step 3: write your code!

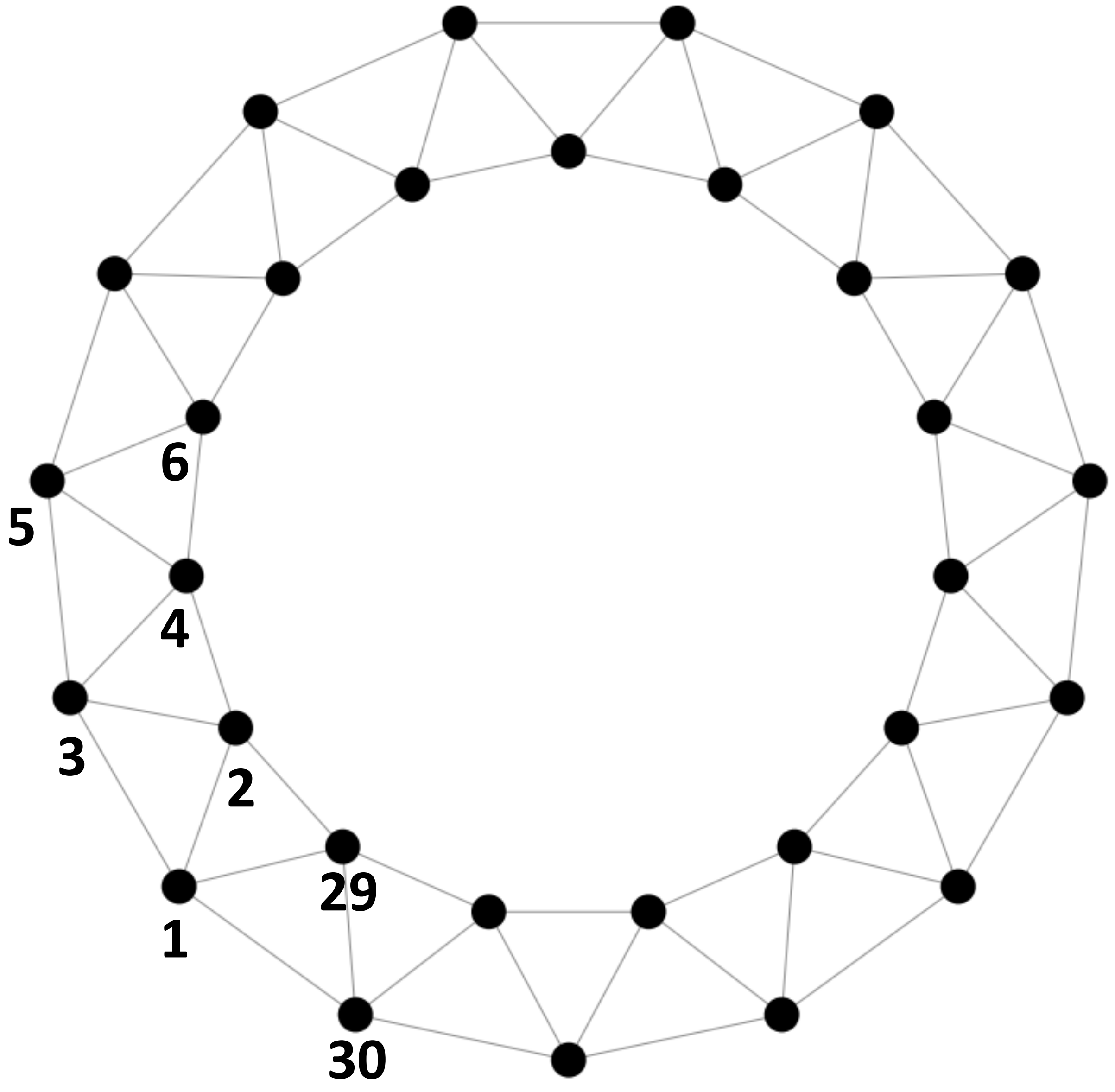
Case study: complex contagions (Centola & Macy, 2007)

<u>Common assumption:</u> small-world networks increase the speed of diffusion for "whatever is to be diffused"

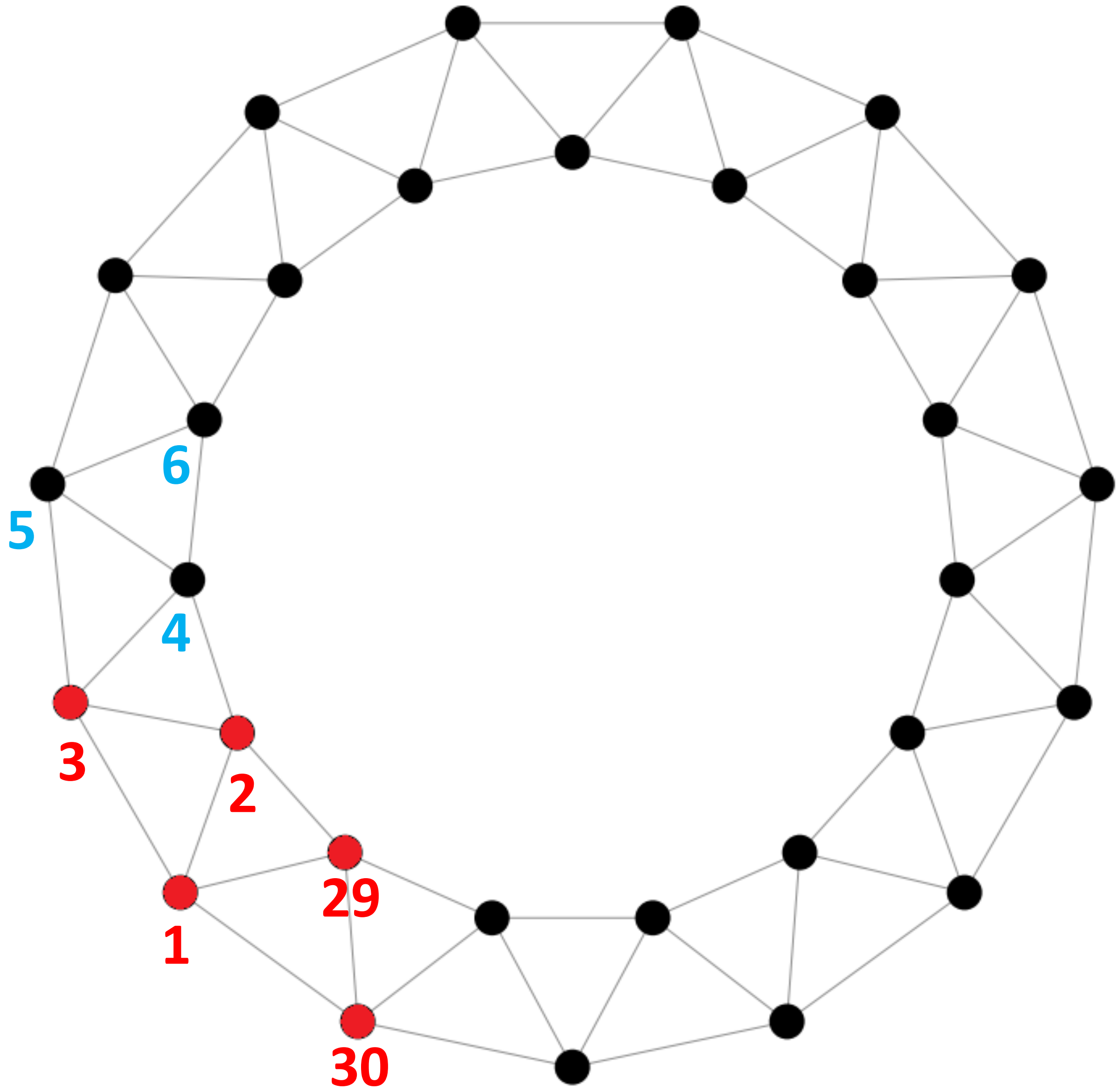
<u>Question</u>

- Sufficiency: Is decreased diameter sufficient to increase speed of diffusion?
- Hypothesis-testing: What is effect of network structure on complex contagions? *(Or: what is the effect of social reinforcement…?)*

Using the Model



Using the Model



29

30

1

2

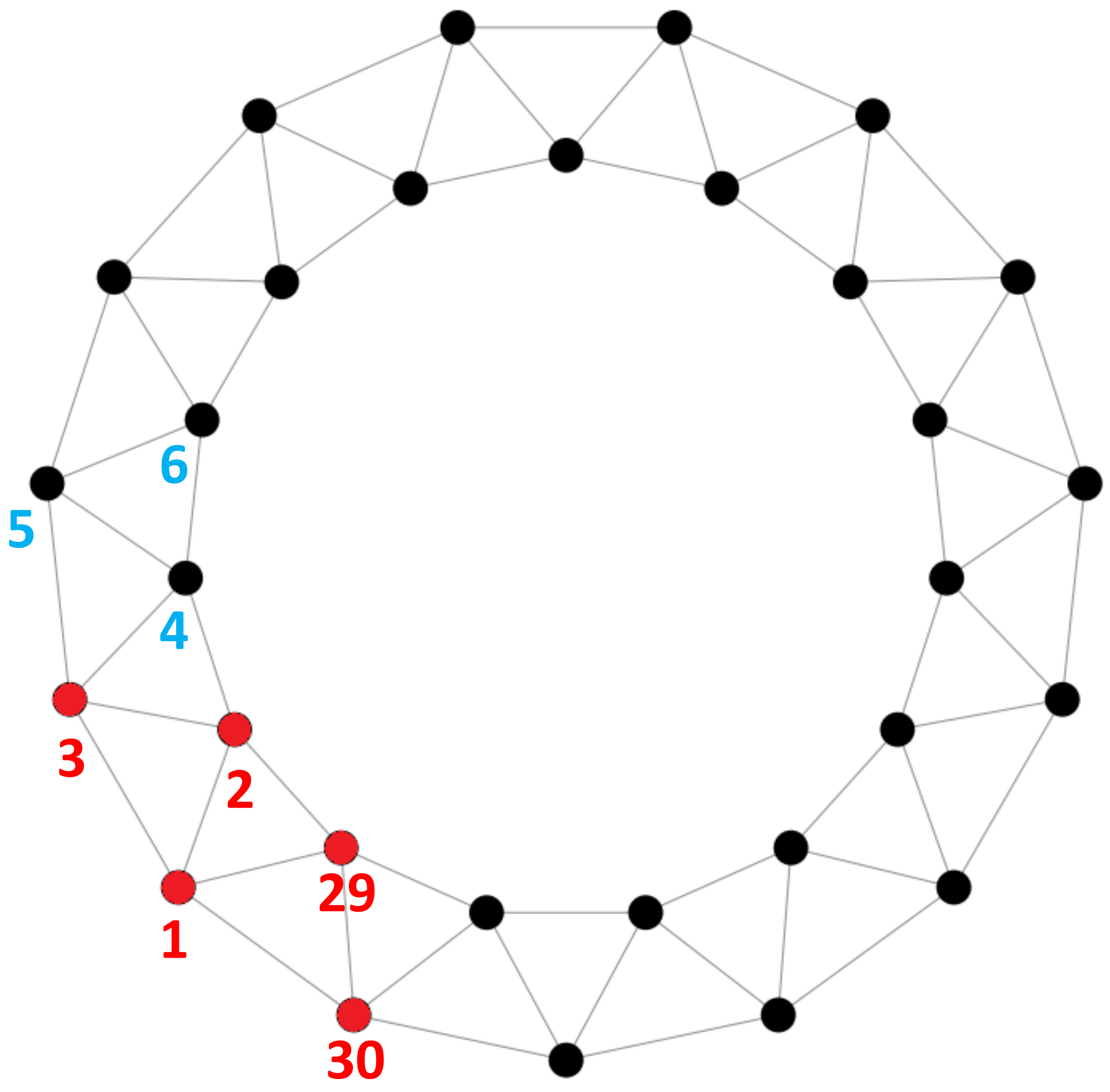
3

4

5

6

Node 4’s neighborhood



29

30

1

2

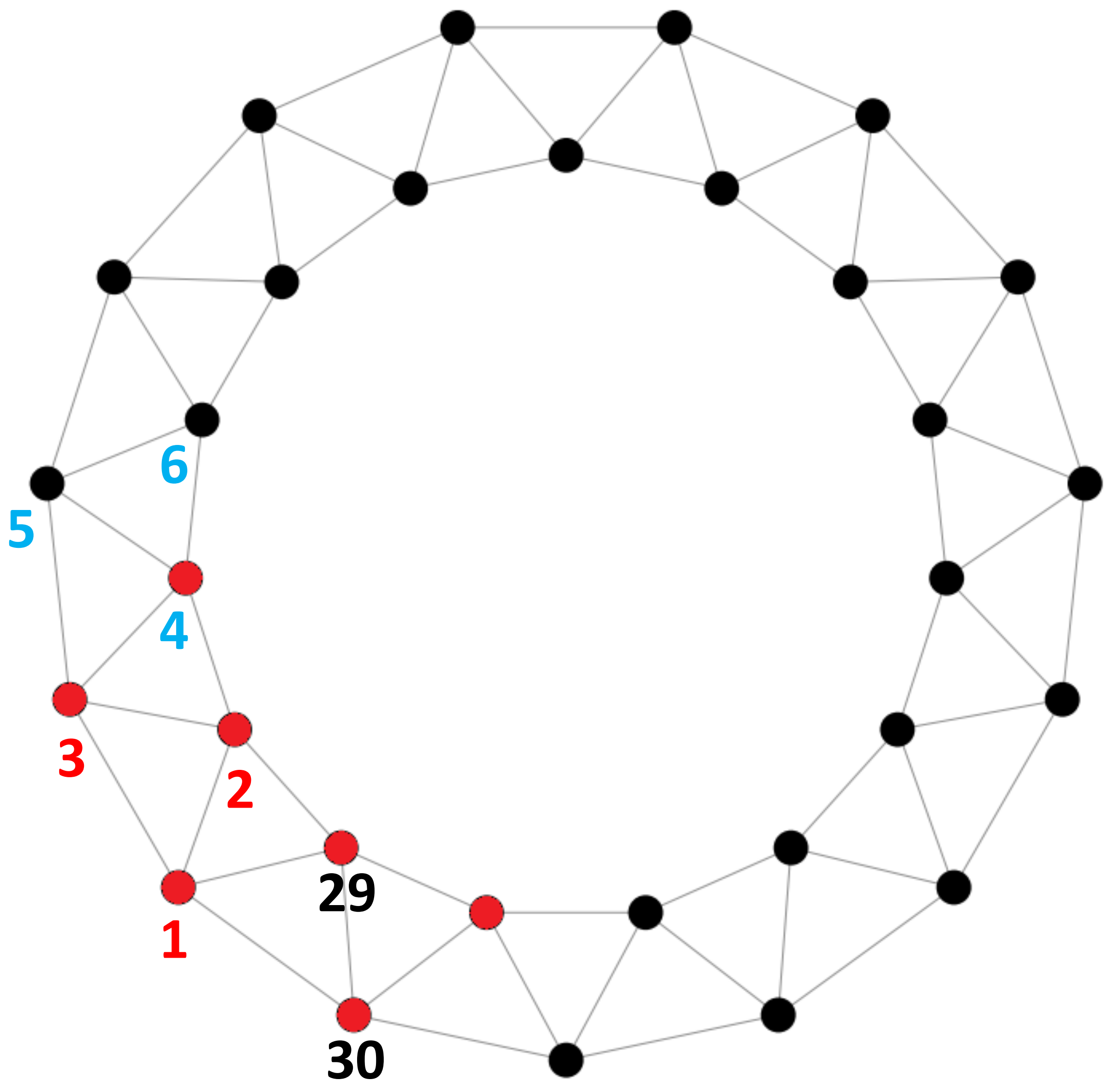
3

4

5

6

Node 4’s neighborhood



29

30

1

2

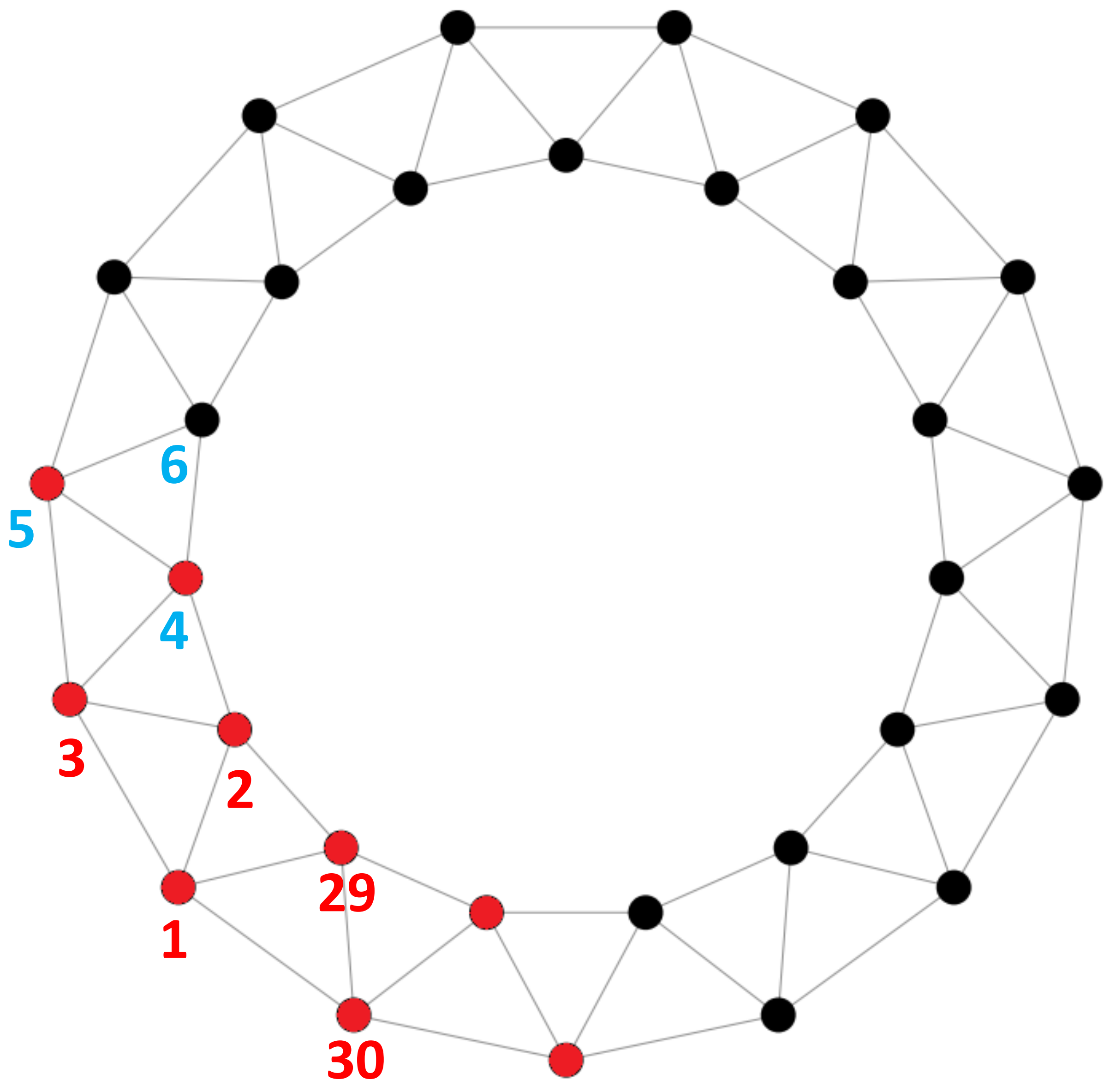
3

4

5

6

Node 4’s
neighborhood



29

30

1

2

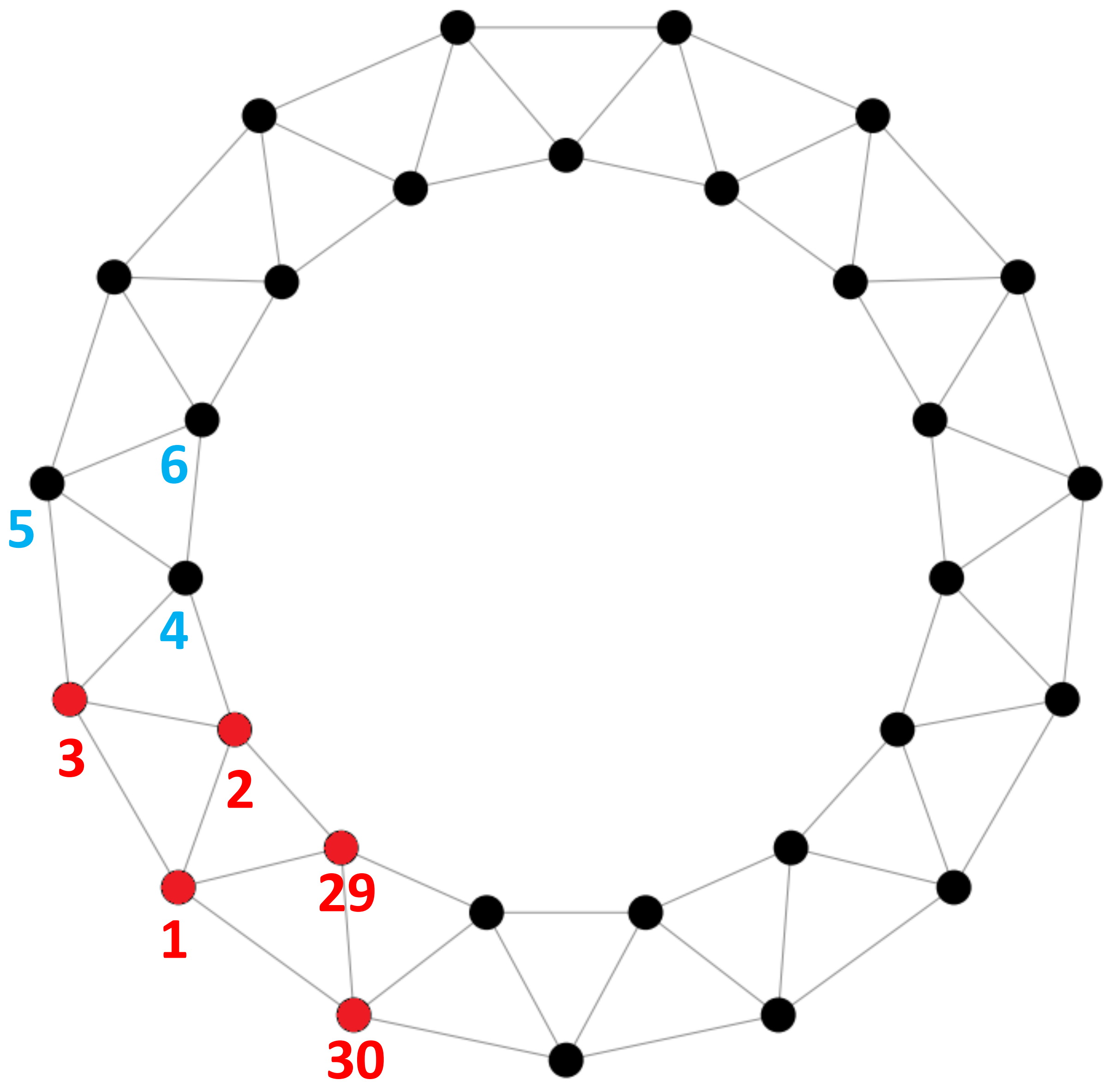
3

4

5

6

Node 4’s neighborhood



29

30

1

2

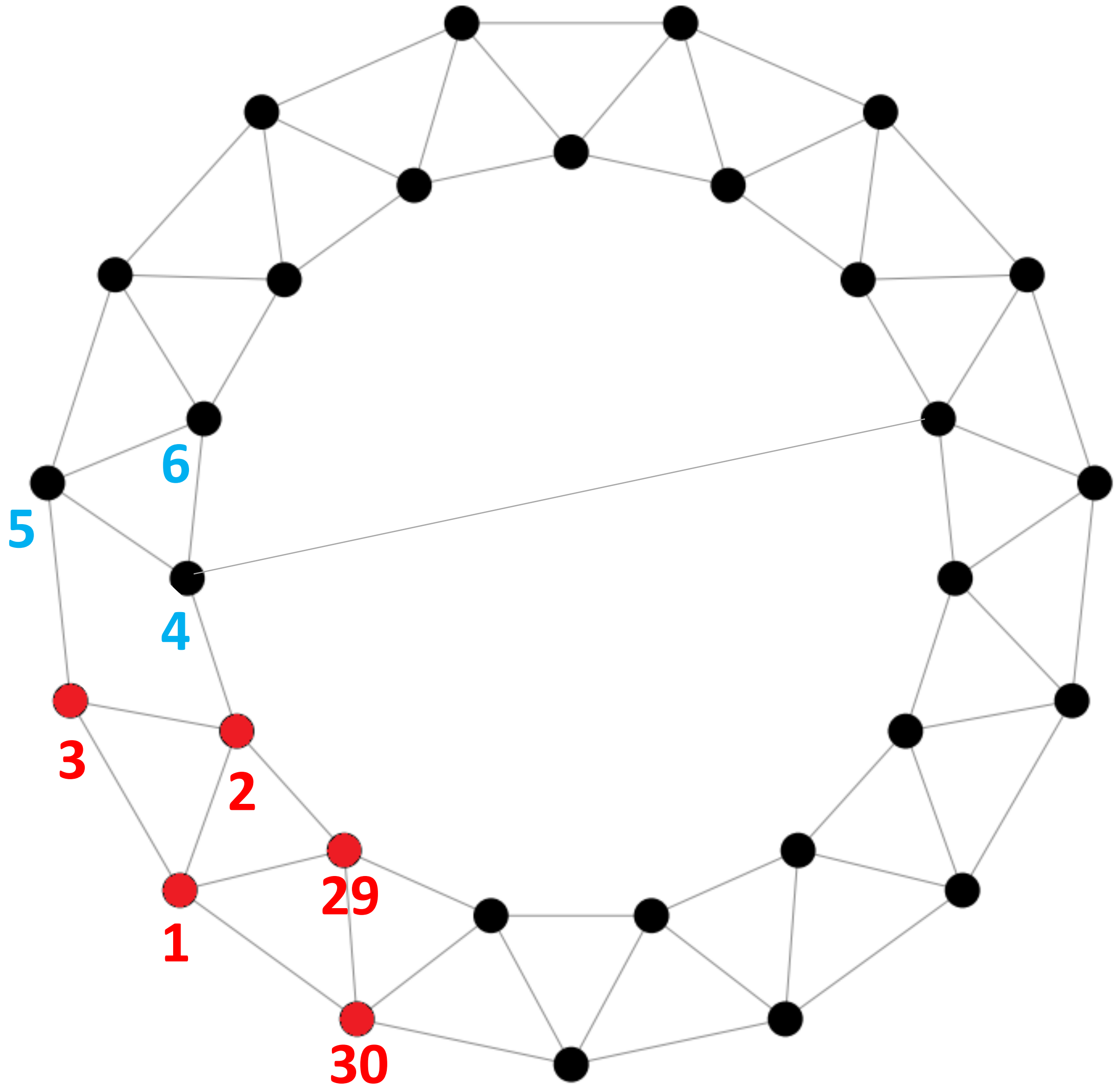
3

4

5

6

Node 4's neighborhood



Code for this model

- Examples in R and Netlogo: https://compsocialscience.github.io/summer-institute/2018/chicago/models

- R example: https://github.com/joshua-a-becker/degroot-simulation

- For faster simulations: use Java

Code for this model

https://github.com/compsocialscience/summer-institute/blob/master/2018/chicago/complex_contagions.R

```
library(igraph)
g = graph.lattice(dim=1, nei=2, length=20, circular=T)
layout = layout.auto(g)
V(g)$state = 0
V(g)[1:5]$state = 1
plot(g, vertex.color=V(g)$state, layout=layout)

V(g)$newstate=V(g)$state
for(i in 1:vcount(g)) {
  neighbor_adoption = mean(V(g)[nei(i)]$state)
  V(g)[i]$newstate = ifelse(neighbor_adoption>=0.5, 1 ,0)
}
V(g)$state = V(g)$newstate
plot(g, vertex.color=V(g)$newstate, layout=layout)
```

Using Formal Models to Design Experiments

Case study: Network Structure of Innovation

The Model:

Lazer, D., & Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, *52*(4), 667-694.

The Experiment:

Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, *109*(3), 764-769.

The Model ("NK Space")

- Problems represented as N bits

- Value of each bit depends on K other bits

The Model ("NK Space")

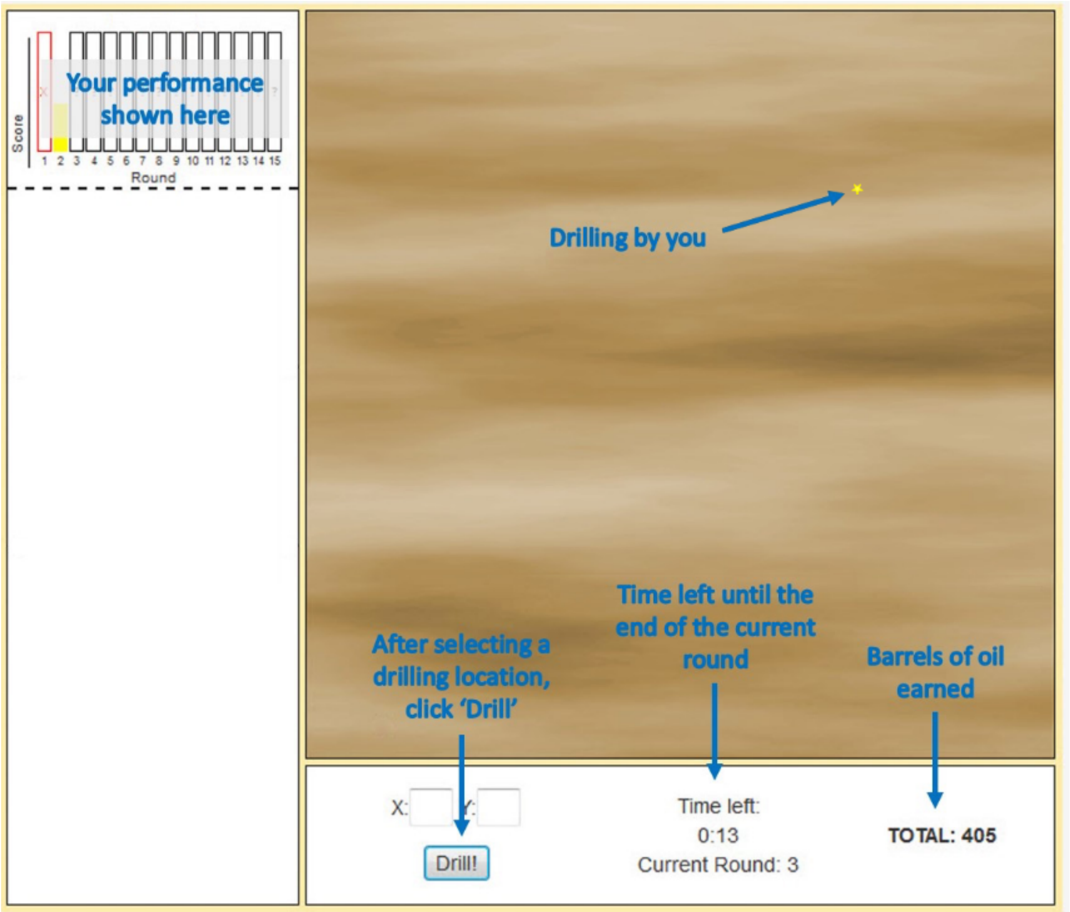
- Problems represented as N bits

- Value of each bit depends on K other bits

- Costly to change multiple dimensions

- Complex interdependence, "rough" landscape

The Game ("Wildcat wells")



Pay attention to scope conditions.

Using a Model to Designing an Experiment

1. Generate reasonable expectations for individual user behavior

2. Use your model: identify conditions most likely to yield results

3. Use your model: generate power tests to estimate sample sizes

Web Experiments Using Free and Open Source Platforms

Components for Web Experiments

- Source of subject recruitment
- Mechanism for assigning subjects to conditions
- Platform for running experimental trial

Experimental Process Flow

1.

Mturk enrollment “HIT”

2.

Subjects contacted
for participation

3.

Invitation sent for
experimental trial

4.

subjects arrive
to “waiting room”

5.

subjects randomized
to condition

6.

Subjects receive
treatment

Options for Single-Subject Experiments

- Amazon Turk – limited experimenter features
- Qualtrix – designed for survey experiments
- jsPsych – optimized for lab style psych experiments

Full Featured (Free & Open Source) Platforms

- TurkServer – javascript based framework, mostly user management
- Otree –python based framework (full featured)
- Empirica.ly – javascript based platform (full featured)

MORE RESOURCES:

https://www.joshua-becker.com/resources