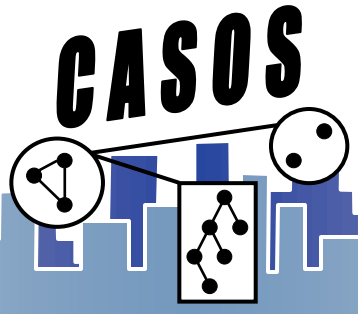


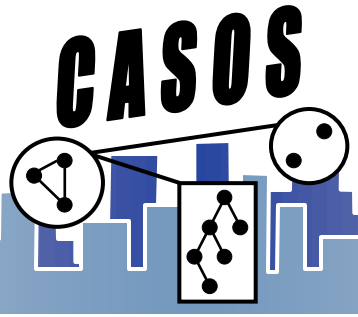
Extended Construct and Simulation Validation

Geoffrey P. Morgan, Michael Lanham
Kathleen M. Carley

institute for SOFTWARE RESEARCH

Carnegie Mellon

Center for Computational Analysis of Social and Organizational Systems
http://www.casos.cs.cmu.edu/



CONSTRUCT MECHANISMS

Agenda

- Definitions
- Mechanics
- Examples
- Discussions

Definitions (1 of 5)

- Construct is a **turn** and **agent-based simulation** tool, useful for **modeling information** and **belief diffusion**.

- Agent-Based Simulation: A simulation of the interactions and attributes of 'autonomous' entities.
  - Agents have **agency** – they take actions that affect their world or other agents.
  - Agents usually have **bounded rationality** – they do the best thing they can do with their current understanding of the world.

- Given that an agent should have agency and bounded rationality, can you think up some examples of agents?

Definitions (2 of 5)

- Construct is a **turn** and **agent-based simulation** tool, useful for **modeling information** and **belief diffusion**.

- Turn-Based Simulation: A type of agent-based simulation, each agent is given a turn to act. Once every agent has been given a turn, a new round of turns begins.
  - The order of agents in the turn may be **static** (fixed at the beginning) or **dynamic** (where it changes every round). The order is dynamic in Construct – to avoid primacy effects.
  - A turn may have multiple phases, what happens in each turn is based on the simulation.
  - Agents may not participate in every phase of every turn, they participate in the phases that apply to their situation.

- Can you think of anything that fits this definition of turn-based?

Definitions (3 of 5)

- Construct is a **turn** and **agent-based simulation** tool, useful for **modeling information** and **belief diffusion**.

- To Model: To represent something in or of the world, removing those details not pertinent to the reason to model.
  - Models can be **physical**, where they represent an object or class of object (e.g., a scale model of a Spanish Galleon), or **abstract**, where they usually represent a process (e.g., recipes are process models describing how food is made).
  - Models must remove details, or they are not models – they are the thing itself. The choice of what details to remove is an art, not a science.

- Besides the examples given, what are some examples of models?

Definitions (4 of 5)

- Construct is a **turn** and **agent-based simulation** tool, useful for **modeling <u>information</u>** and **belief <u>diffusion</u>**.

- Information Diffusion: The process by which knowledge moves through a social group
  - Knowledge can be of varying "sizes" – but the "size per bit" should be consistent in each simulation. "James was seen with Sally at Seviche" can be a knowledge bit, as can "F-22 Pilot Operations", but they should not be the same number of bits inside the same simulation.
  - Social Groups are defined by the networks of interacting actors. This makes the simulation **network-centric**.

- What phenomena do you think can be modeled with information diffusion processes?

Definitions (5 of 5)

- Construct is a **turn** and **agent-based simulation** tool, useful for **modeling information** and **belief diffusion**.

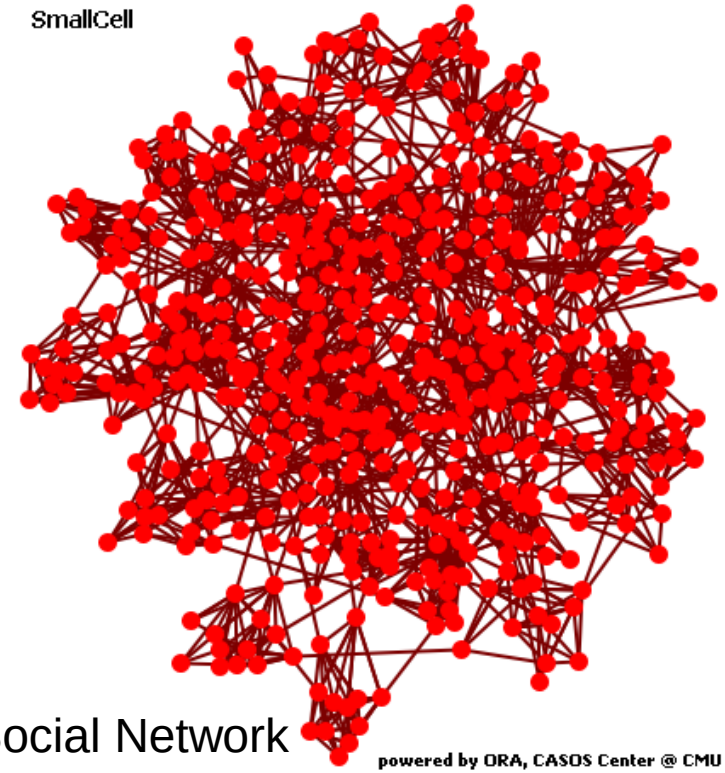
- Belief Diffusion: The change in beliefs of actors in a social group over time.
  - Beliefs cannot be evaluated for truth.
  - Knowledge can contribute to or deny a belief.
    - Belief: "Cats are better house-pets for a family than dogs."
    - Supporting Evidence: "Cats live longer than most breeds of dog."
    - Contrary Evidence: "Cats are less affectionate than dogs."

- What other kinds of beliefs might be interesting to model using this kind of simulation?

Mechanics (1 of 6)

- Construct is **agent-based** and **network-centric**

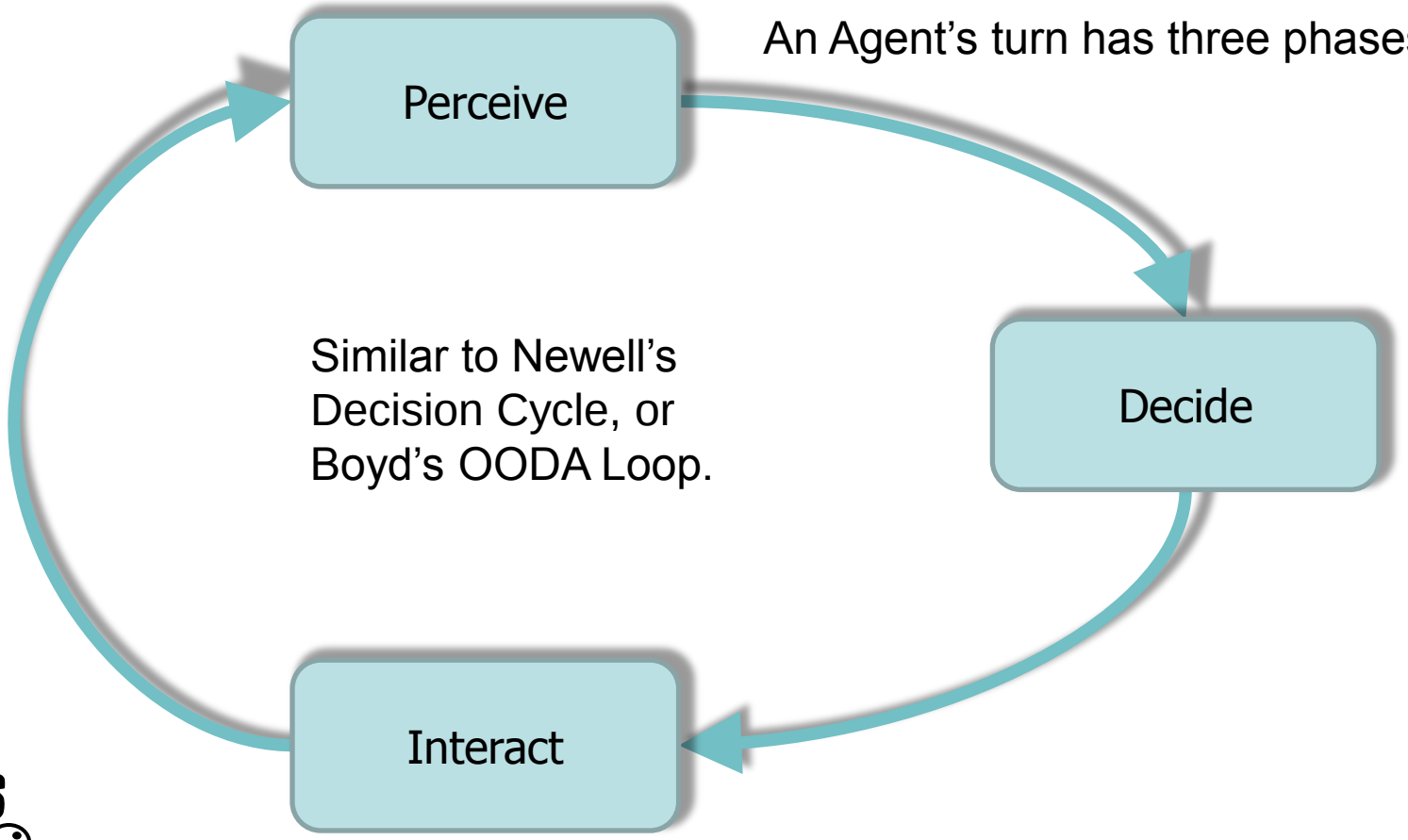
Agent



Social Network

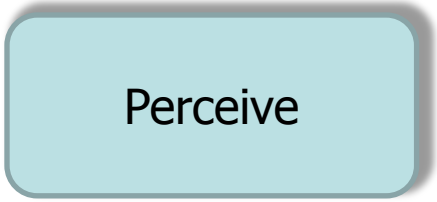
Mechanics (2 of 6)

- Construct is a **turn-based simulation**.



Mechanics (3 of 6)

- Construct agents have **bounded rationality**.



The perception of other actor's knowledge is an error-prone process. Agents may misunderstand what other agents tell them, make incorrect inferences, or simply remember things incorrectly.

These errors are important and useful because:

a) Real humans (and other kinds of agents) make these kinds of errors

b) These errors inform future interaction decisions by the agent.

Bounded Rationality: Group schemas inform alter perceptions

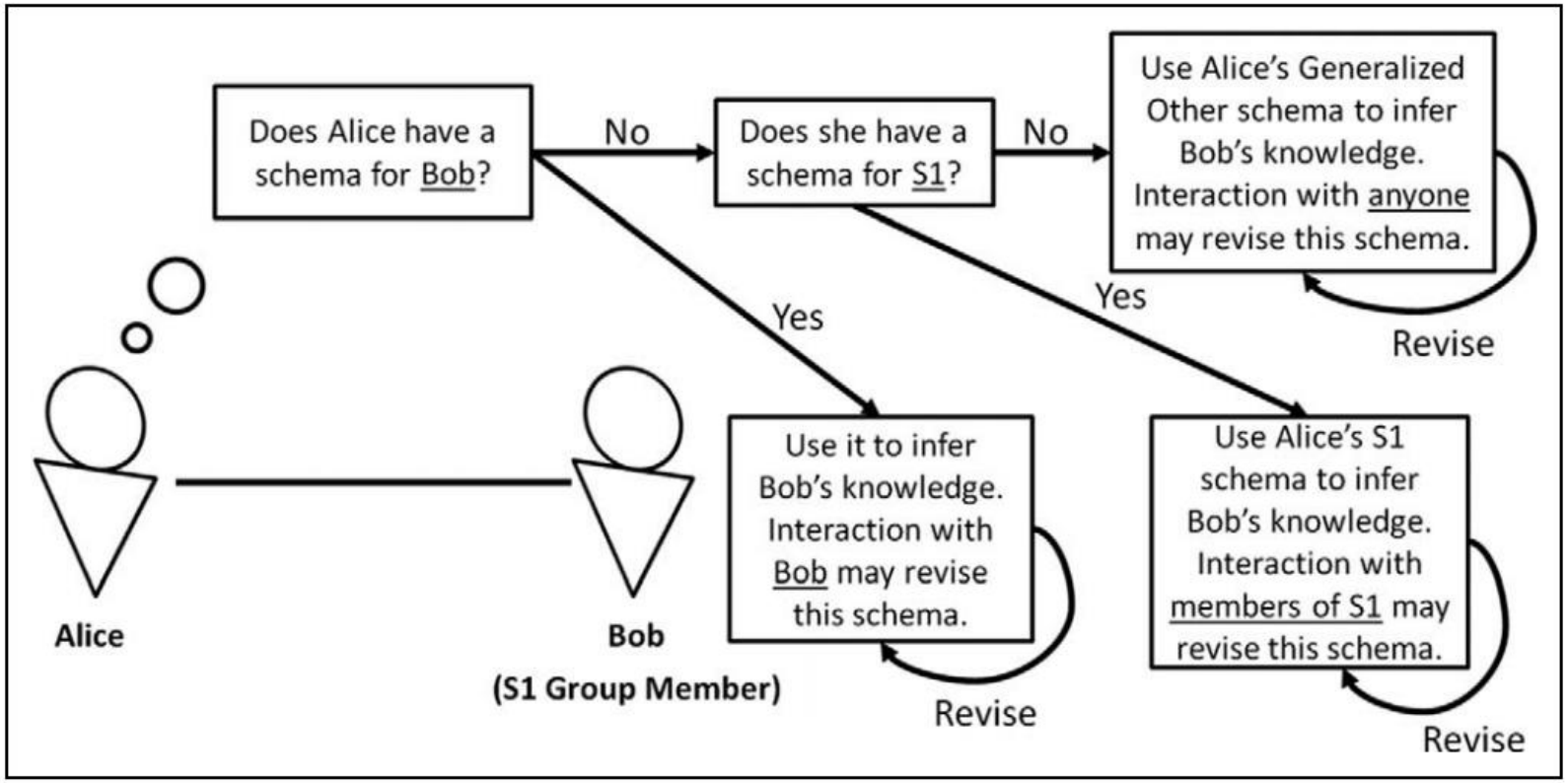
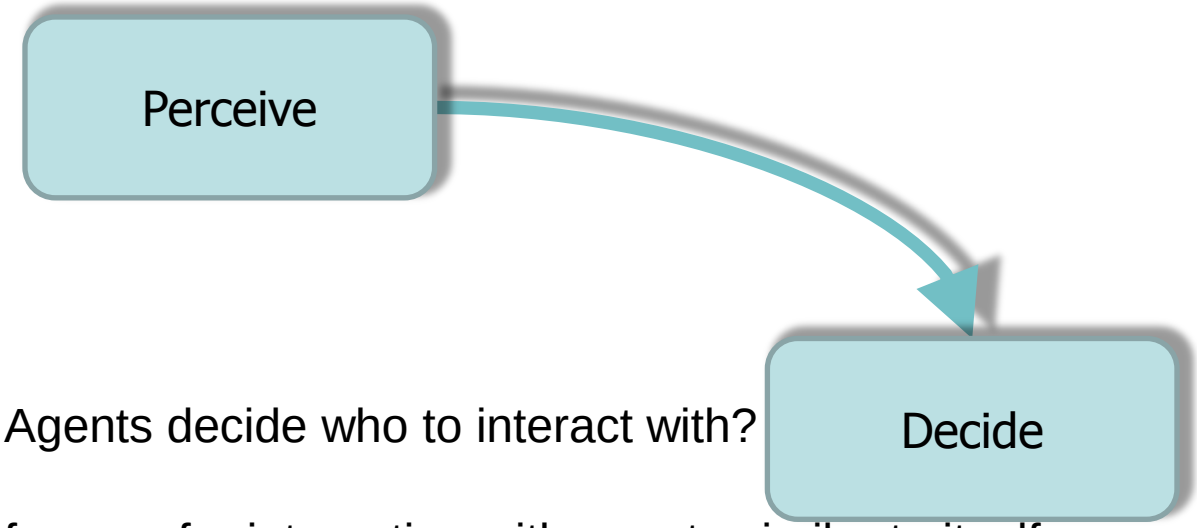


Figure 1. Alice needs to infer Bob's knowledge to determine her likelihood of interacting with him.

Joseph, K., Morgan, G. P., Martin, M. K., & Carley, K. M. (2013). On the Coevolution of Stereotype, Culture, and Social Relationships: An Agent-Based Model. *Social Science Computer Review*, 0894439313511388.

Mechanics (3 of 6)

- Construct agents have **bounded rationality**.



How do Construct Agents decide who to interact with?

- Homophily: A preference for interacting with agents similar to itself.
- Expertise-Seeking: A preference for interacting with agents with rare knowledge.

These two drives are informed by the perception process.

Note. Construct also supports propinquity (adjacency) as a moderator of interaction for those spatially inclined.

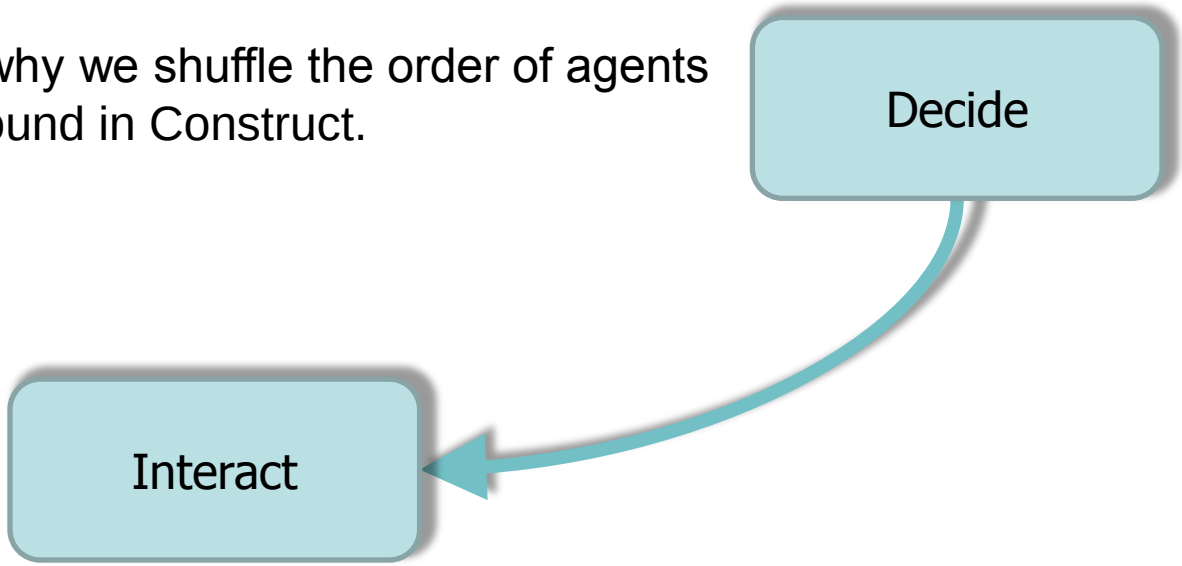
Mechanics (4 of 6)

- Construct agents have **agency**.

The decision of who to interact with is informed by the agent's own perceptions of the world and the agent makes that decision independently.

Interacting with an agent can block others from interacting with that agent – thus, the actions of agents can have important consequences for others.

It's also why we shuffle the order of agents per round in Construct.

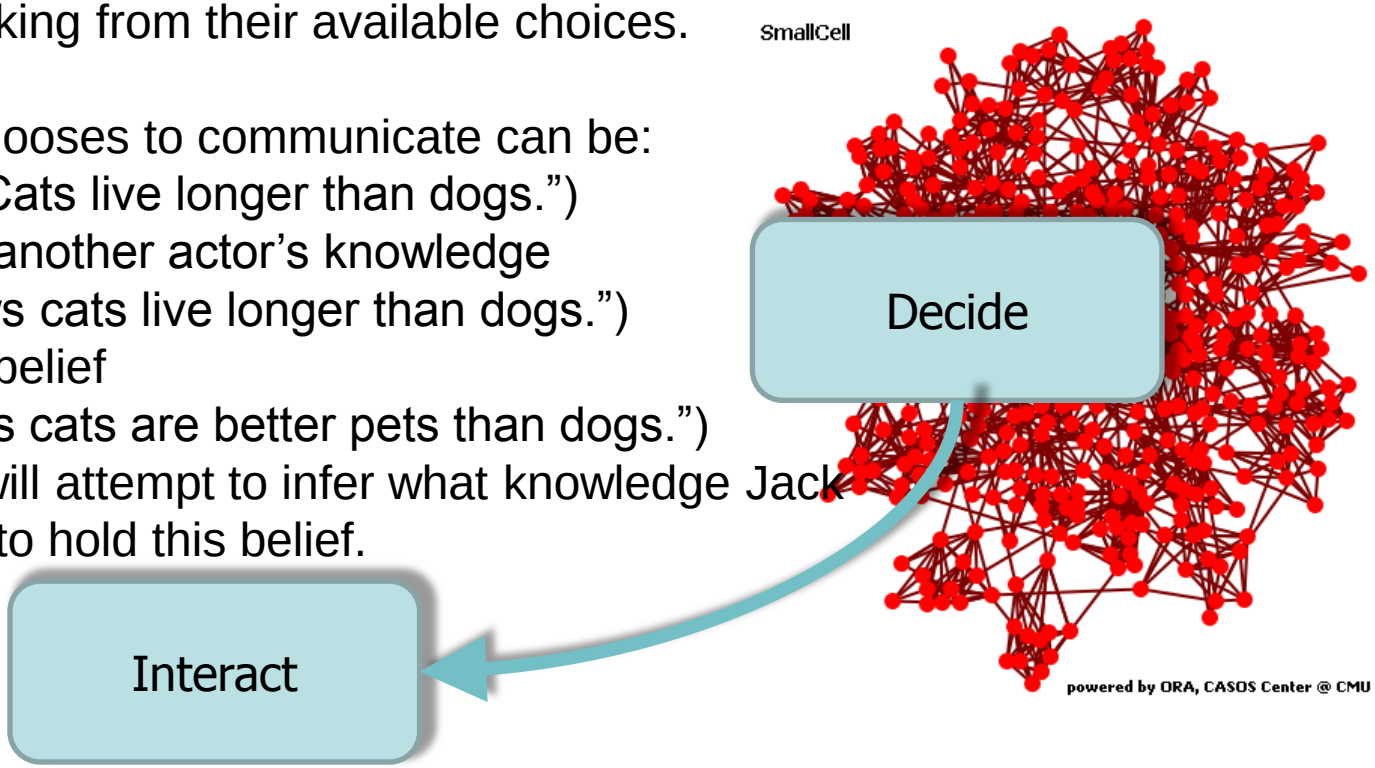


Mechanics (5 of 6)

- Construct is used to model **information diffusion**.

Agents are usually constrained in who they can interact *with* based on their social network. They make the best choice based on homophily and expertise-seeking from their available choices.

What the agent chooses to communicate can be:
- Knowledge ("Cats live longer than dogs.")
- Perception of another actor's knowledge ("Jack knows cats live longer than dogs.")
- Perception of belief ("Jack thinks cats are better pets than dogs.") The agent will attempt to infer what knowledge Jack might have to hold this belief.



Mechanics (6 of 6)

- Construct is used to model **belief diffusion**.

Fact 7: Cats damage furniture.
Fact 8: Cats are passive aggressive.
Fact 9: Cats leave trophies.
=> I don't want a cat!

B: Cats are better pets than dogs.

3 2 1

Decide

Interact

Method

- To be able to use Construct successfully, you need to be able to define:
  - Agents – who/what are the decision-makers?
  - Networks – How is interaction between these agents constrained?
    - This can be provided by a stylized network generator if interested in more abstract questions
    - Or can be provided by data if interested in a specific problem
- You also should know:
  - Metrics – what do you want to measure?
  - Experimentation – what are you going to compare?

More Information

- For the Summer Institute, we called Construct from ORA.
- Construct can also be run as a stand-alone tool.

- A lot more information can be found on the Construct tool page at CASOS

http://www.casos.cs.cmu.edu/projects/construct/index.php

VALIDATION

Definitions

- Validation – a set of techniques for determining whether or not a model is valid. Used for both internal validity, matching with other models, and matching with non-computational data

- Calibration – a set of techniques for tuning a model to fit detailed non-computational data.
- Training – procedures for supplying feedback to computational models that can learn.
- Verification – A set of techniques to confirm that a validated model is appropriate to a given specific context

- Docking – a set of techniques for determining the level of comparability or equivalence of two models, can sometimes be used to inherit 'validation' of a prior model

Traditional Validation Approaches

- Requirements engineering and formal methods
- Evolutionary verification and validation
- Statistical methods alone
- Expert systems (not usually done because simulations are usually numerical)
- Domain experts (human experts/subject matter experts)
- Response Surface Methodology (validation against empirical data)
- Docking, including against math and system dynamics models

Validation is Difficult

- Models are a subset of reality; model assumptions may not match the reality
- Cognitive bias of human modelers/validators
- Validation is knowledge intensive
- Complexity and stochasticity of social agents
- Agent history (non-Markovian), starting conditions, etc.
- Validation consumes a significant amount of time and resources
- Quality and quantity of empirical data vary

Least developed area of Multi-Agent Social-Network computational modeling

Methods and Levels for Validating a Computational Model

- Validation techniques vary in
  - Method
  - Level
  - Intensity
  - Purpose

Aspects of Validation for Simulation

- Internal validity - error free code
- Parameter validity - parameters match
- Process validity - processes fits
- Face validity - right type of things
- Pattern validity - pattern matches observed
- Value validity - values match
- Theoretical validity - theory fits

A Caveat

- Computational modeling is sufficiently complex that a single individual in a single research period (e.g. 6 months to a year) can not build, analyze, **and validate** a computational model.
- Most models take multi-person years to build and analyze.
- Data collection and analysis from a virtual experiment often takes as long as a human experiment and requires statistical training comparable to that required for human experiments.

Face Validity

- Is the model a reasonable simplification of reality?
- Techniques to increase face validity:
  - Set parameters based on real data
  - Model a specific organizational or inter-organizational process
  - Show that others have made similar assumptions
  - Discuss model limits and how left out factors may or may not affect results
  - Don't over-claim model applicability

Stylized Fact Validation

- Simple techniques for seeing of model results are reasonable.
- Techniques to demonstrate validity:
  - Are there stereotypical facts about the problem that this model generates; E.G., Models of organizational evolution should predict liability of newness.
  - Are there behaviors that any model of this ilk should generate; E.G., All diffusion models should generate an s-shaped adoption curve, all neural networks should take a long time to train.
- These are non-surprising findings but if model can't generate them then it is not valid.

Stylized Fact Validation and Construct

Stylized Fact Validation, Definition

- Stylized Fact Validation allows us to leverage insights gathered about a phenomena of interest to evaluate a computational model
  - Generally, the model should not explicitly implement the insights, but instead present a theoretical model that allows these insights expression.
  - Insights can be drawn from the broad sample of research, often "Citation Classics" are used to establish the model's validity.

Kaldor (1961) defined this technique with his six "facts", which were statements on economic growth.

Relevant Research (Construct example)

- Diffusion of new information follows an S-Shape Curve (Rogers, 2010)
- Improvement in task competency in cliquish groups will have more marginal variation over time. (West, Baron, Dowsett, & Newton, 1999)
- Groups with some heterogeneity outperform purely homophilous groups (Ancona & Caldwell, 1992)
- Heterogeneous groups are more likely to discover novel information from outside the group. (Granovetter, 1983)
- Individuals are more likely to interact in-group than out-group. (Blau, 1977)

Construct Extension: Adding groups to Construct

- To make Construct more useful and more scalable, we've added reasoning about groups to Construct. In the latest version of Construct, agents:
  - Have an impression of what all other agents know
  - Have an impression of what some groups know
  - Have an impression of what specific alters know

- We use Mead's Generalized Other (Mead, 1925), and Wegner's Transactive Memory (1987) as the theoretical framing.

Group-Based Insights?

- What aspects of group reasoning do you think the model should include?

Selected Insights

- Group affiliation can trigger a tendency to favor one's own group at the expense of others. (Tajfel, Billig, Bundy, & Flament, 1971)
- Group perceptions moderate our behavior. (Mead, 1925)
- Our perceptions of others are often based on expected roles, social norms, and categorizations (Greenwald & Banaji, 1995; Heise, 1979)