Carnegie Mellon
institute for SOFTWARE RESEARCH

Agent-Based Simulation and Meta (Multi)-Modeling

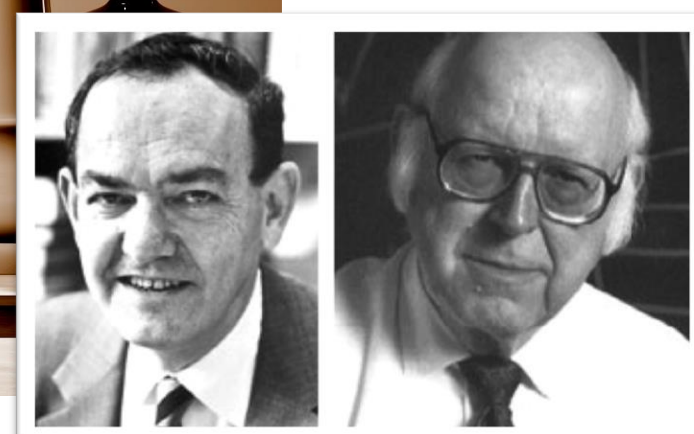
Geoffrey P. Morgan

Tutorial: SBP-BRiMS 2017

CASOS

INTRODUCTION TO AGENT-BASED MODELING

The Origin Story (Apocryphal)



Herb Simon and Allen Newell

Modeling

- 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.

Definition: Agent-Based Simulation

- 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.

Bounded Rationality

- Agents have **bounded rationality**.

Perceive

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.

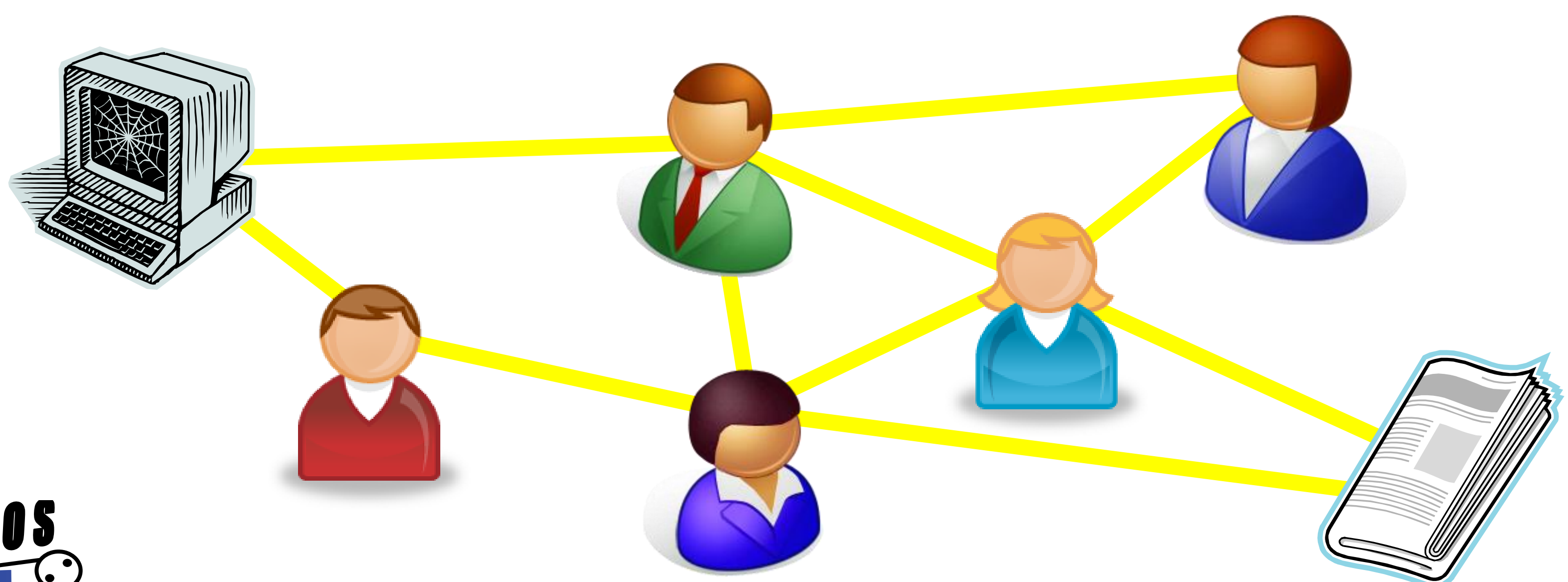
These errors are important and useful because:

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

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

Agents Can Be Human or Non-Human

- Agents are often human but do not have to be
  - the type of each agent in the simulation can be specified
  - often, non-humans are represented as different types of agents
  - most non-human agents cannot initiate communication, but must wait for a human to contact them



When do you build and agent based model?

- When there is a natural representation of things as agents
  - When there are decisions and behaviors that can be defined discretely (with boundaries)
  - When it is important that agents adapt and change their behavior
  - When it is important that agents learn and engage in dynamic strategic behavior
  - When it is important that agents have a dynamic relationships with other agents, and agent relationships form and dissolve
  - When it is important that agents form organizations and adaptation and learning are important at the organization level
  - When it is important that agents have a spatial component to their behaviors and interactions
- When the past does not seem to predict the future
- When scaling-up to arbitrary levels is important
- When structural change of process needs to be a result of the model, rather than an input to the model

The Many Faces of Computational Models

- Models as theory
  - The model does the task it seeks to explain
  - Social Turing test
- Models as agents
  - Tournament participants
  - Substituting for humans in the lab
- Models as virtual world
  - Social theory from the ground up
  - A-life
- Models as empirically grounded theory
  - Empirical data as integral part of model
  - May be ethnographic
- Models as hypothesis generators
  - Response surface analysis
  - Systematic generation of hypotheses

Why Use Computational Modeling and Analysis?

- *Ethical:* Cannot test policies on real populations
- *Preparatory:* Can create hypothetical situations with more potency than existing ones – Can examine wide range of scenarios – Enabling systematic imaginative thinking
- *Cost effective:* Creating new technologies, procedures and legislation for data collection is expensive
- *Faster:* Real time evaluation of existing systems is too time consuming
- *Appropriate:* The world and the simulation are complex non-linear dynamic systems
- *Flexible:* Response to novel situations requires rapid evaluation of previously unexamined alternatives

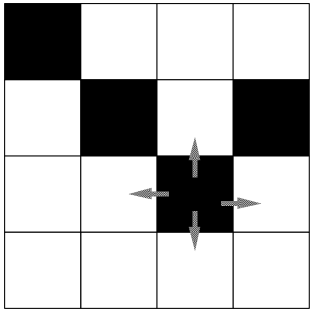
How to Use Computational Models

- Test bed for new ideas
- Predict impact of technology or policy
- Develop theory
- Determine necessity of a posited mechanism
- Decision making aids
- Forecast future directions
- What if training tools
- Suggest critical experiments
- Suggest critical items for surveys
- Suggest relative impact of different variables (factors)
- Suggest limits to statistical tests for non-linear systems
- Substitute for person, group, tool, etc.. in an experiment
- Hypotheses generators

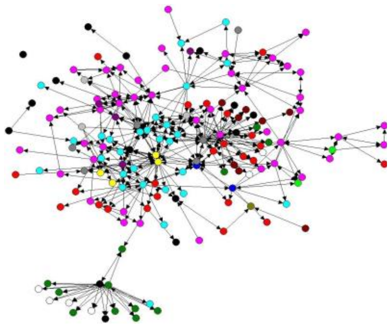
Interaction Topologies

- Agents can move in space (discrete or continuous)
- Cellular automata have agents interacting in local "neighborhoods"
- Agents can be connected by one or more networks
  - static or dynamic
- Agents can move in physical space
  - over Geographical Information Systems (GIS) tilings
  - From one lat-lon to another
- Sometimes *spatial interactions are not important*

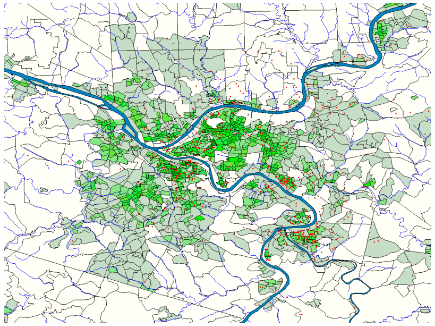
Euclidean



Grid



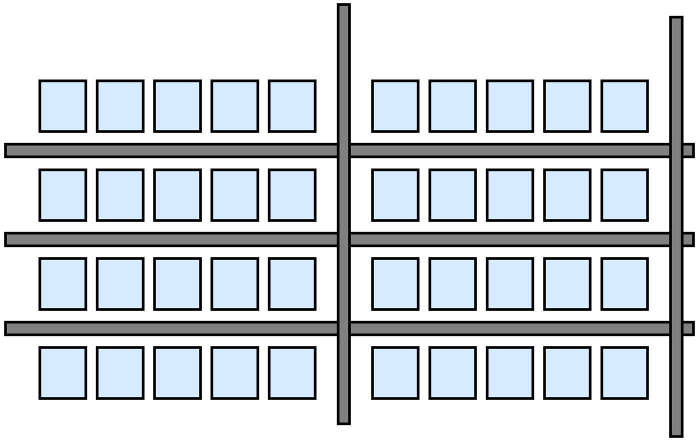
Network



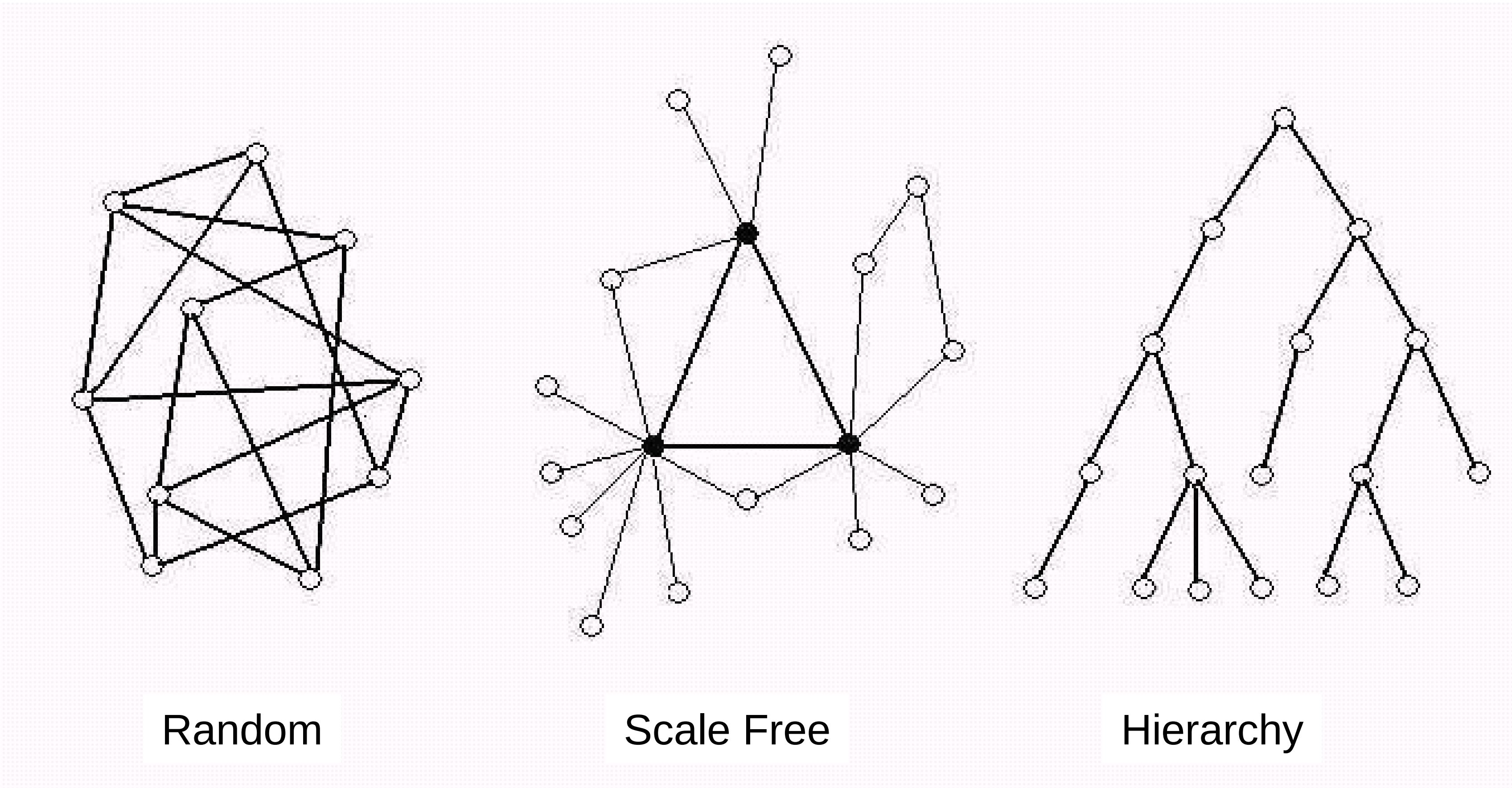
GIS

Latane - City Blocks

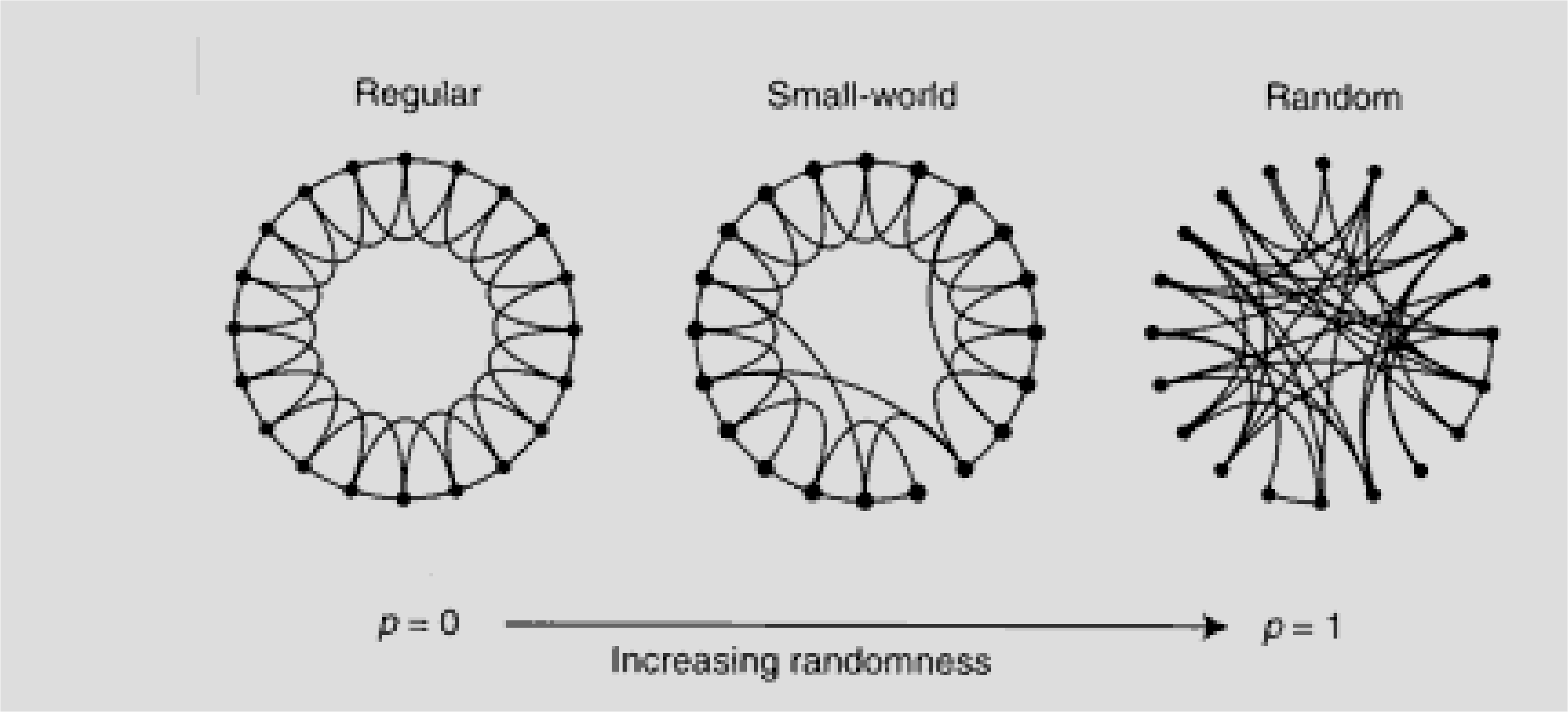
- Cellular Automata
- Opinion formation.
  - Agent has an opinion.
  - Agent's opinion as affected by neighbors.
  - Agent can move or change opinion.
- Movement along grid lines.
- Close off areas as parks, etc.
- Similar to a-life.
  - Distinction is diagonals.



Different types of Networks (Space)



Small World Networks

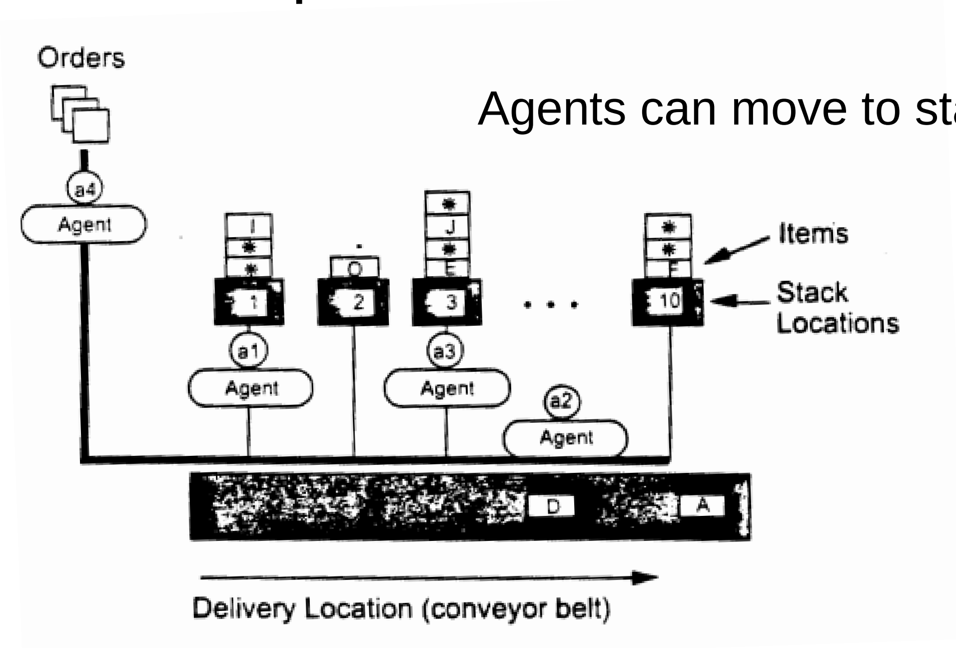


Small-Worlds by Duncan Watts, 1999

EXAMPLE AGENT-BASED MODELS

PLURAL-Soar: Agents in the Warehouse

Model of a small organization, a Warehouse, where each agent, a Soar agent, is responsible for locating and retrieving items as requested in the storehouse.



Agents can move to stacks to find things

If more than one agent is at a stack, they form a line.

Only the first agent in line at a stack can interact with it.

Carley, K.M. & Kjaer-Hansen, J. & Newell, A. & Prietula, M. (1992). Plural-Soar: a Prolegomenon to Artificial Agents and Organizational Behavior. *Artificial Intelligence in Organization and Management Theory*, 87-118. Amsterdam: North-Holland: Elsevier Science Publishers.

The Soar Cognitive Architecture

- Soar used to be an acronym: State Operator And Result
- First development in 1983 as collaboration between John Laird and Allen Newell
- Still going strong! Recently held the 37th Soar Workshop

http://soar.eecs.umich.edu/

- Soar has these key elements:
  - Working Memory Elements (WMEs) which define a state
  - Rules
    - Elaboration Rules – which elaborate the state
    - Proposal Rules – which propose operators (actions)
    - Application Rules – which apply operators to change the state

Soar WME Example: Blocks-World

8 CHAPTER 2. THE SOAR ARCHITECTURE

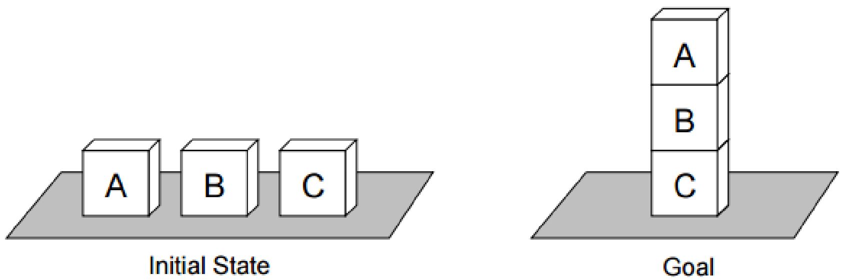


Figure 2.2: The initial state and goal of the "blocks-world" task.

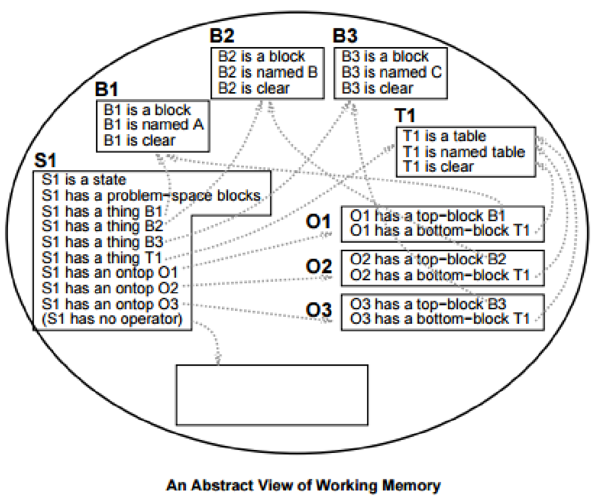
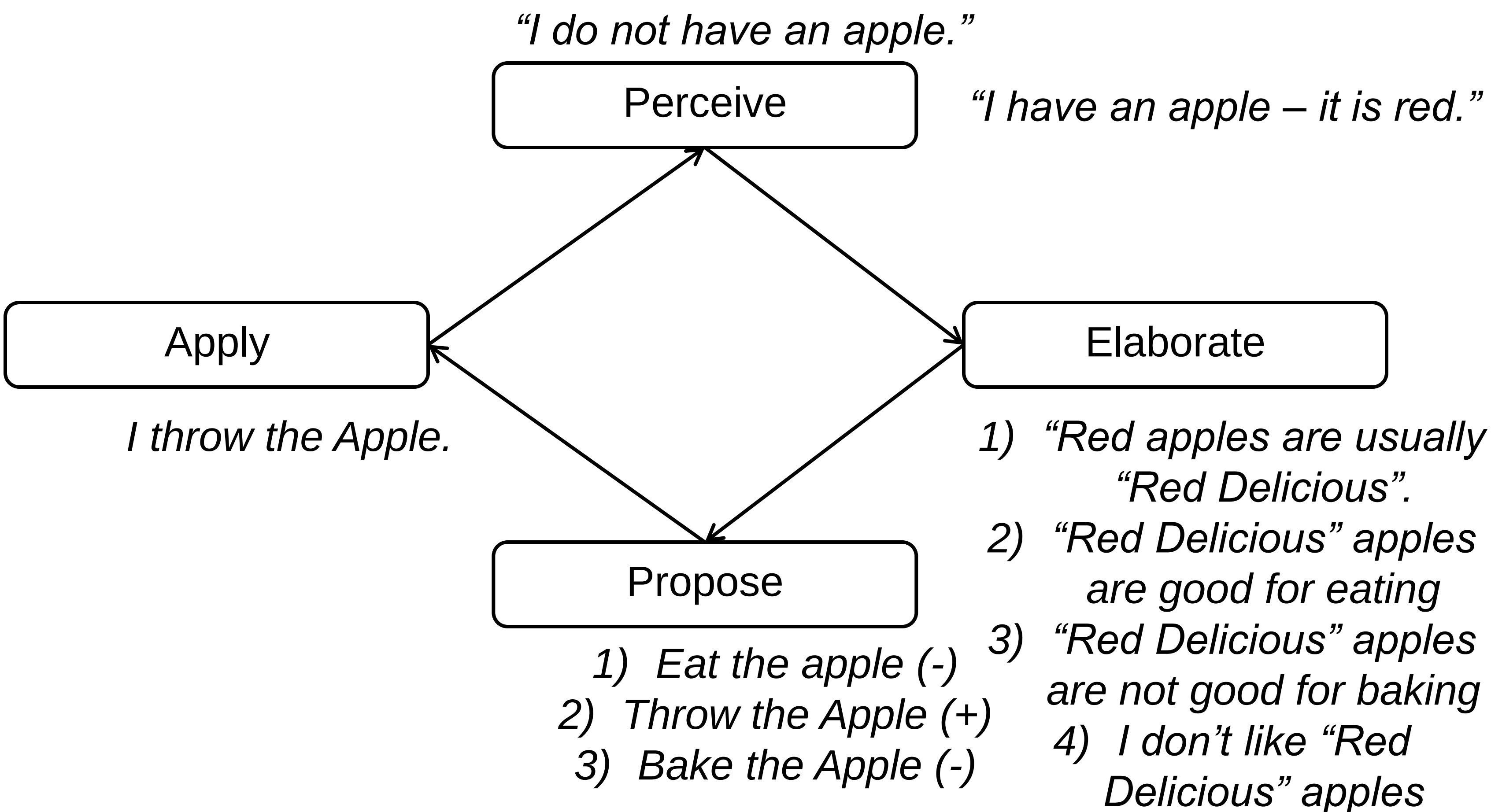


Figure 2.3: An abstract illustration of the initial state of the blocks world as working memory objects. At this stage of problem solving, no operators have been proposed or selected.

From p.8 of
Laird, J. E., & Congdon, C. B. (2015). The Soar User's Manual Version 9.5. 0.

Soar Agent Cycle



Additional Specification of the Warehouse

- 10 stacks
- 3 items on each stack, at first
- 15 orders
  - Each order is for one item
- All items are unique
- Agents work only one order at a time
- No communication channels except talking
- Everyone can hear everything everyone else says

4 Soar Models, One Task

- Agent B (Basic)
  - Can interact with the warehouse
  - Searches for each order item independently
- Agent BL (Basic + Location Memory)
  - Remembers orders as last seen in visited stacks
  - (Remember: Items are moved to get other items often.)
- Agent BC (Basic + Communication)
  - Can ask questions of other agents
  - Remembers other people's question to respond to if the item is seen later
  - Will prefer to ask than to search
- Agent BCL (Basic + Location Memory + Communication)
  - Remembers visited stacks and can ask questions
  - Trusts its own memory over answers to questions

Plural-Soar Experimental Design

Parameter	# of Values	Values
Number of Agents	5	1,2,3,4,5
Agent-Type	4	B, BL, BC, BLC
Outcomes		
Turns to complete task		
Decision Cycles To complete task		
Physical Effort to complete task		
Time waiting to complete task		
Questions asked		
Questions answered		

What are the constants?

Results

- One of the first Soar agents that did a cooperative task based on small-group behavior studies
  - Each agent required its own process
  - The warehouse was represented by a central data file they each interacted with in turn
- Effort, time, and cycles decreased as more agents were involved
- The most questions are asked and answered when the organization is mid-size (2 or 3)

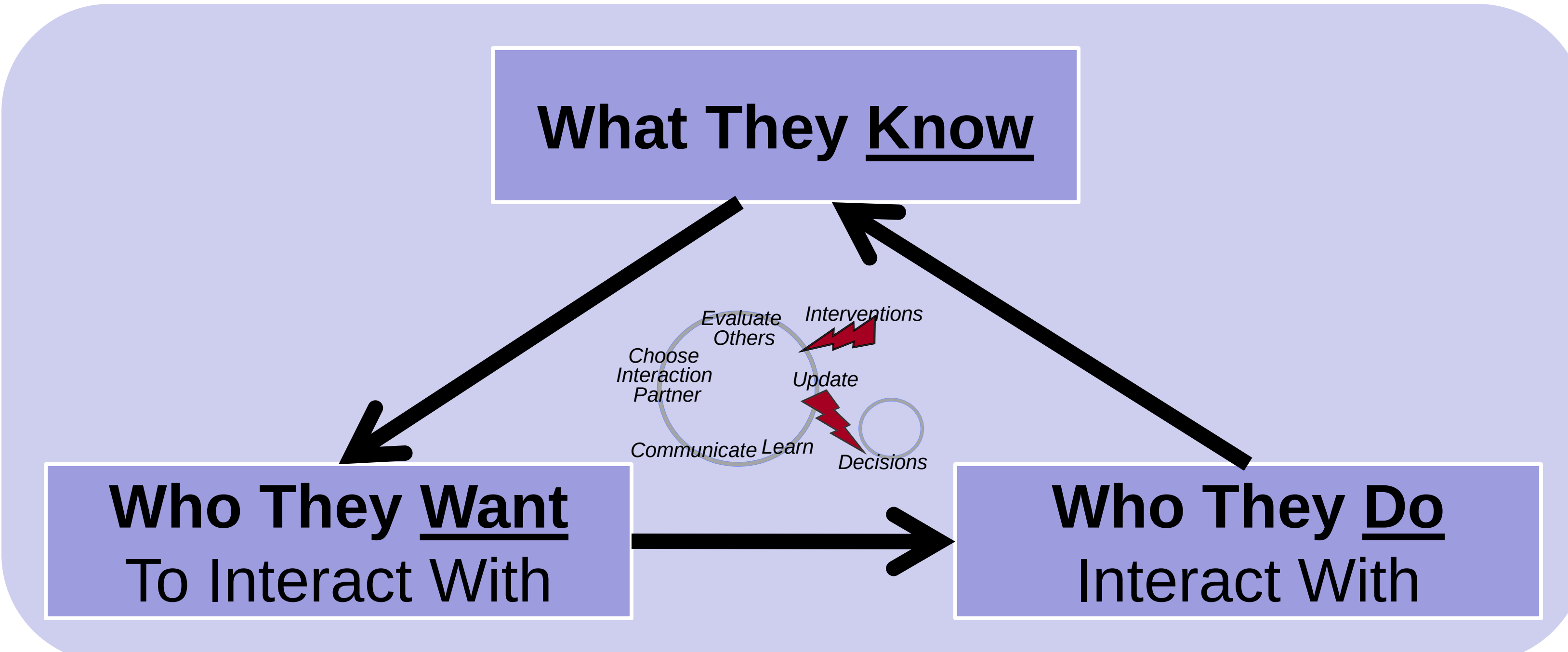
Construct

- Dynamic-Network Agent-Based simulation model for examining information diffusion and social change
- First multi-agent network model in socio-cultural area
- Features
  - Co-evolution of social structure and culture
  - Co-evolution of agents and their societies
  - Co-evolution of social and knowledge networks
  - Agents learn through interaction
  - Agents need not be “people”
  - Multi-fidelity input is possible
    - Exact knowledge network
    - Group level probabilities
- Refactored in 2009 to use modern agent-based techniques
- Multi-Level as of 2015

How Has Construct Been Applied?

- Organizational theory
  - what organizations are efficient, and what ones are not?
  - how robust are certain organizations from threats and attacks?
- Key player forecasting
  - how will a network evolve over time, and what could it look like?
  - what agents are likely to be important due to network position?
- Intervention evaluation
  - how will specific interventions affect certain types of decisions?
  - what are the secondary effects of different interventions?

What is Construct?



- Construct is a sophisticated agent-based dynamic-network simulation engine
  - the agents, social network, and knowledge base are dynamic
  - the effects studied are complex, varied, and highly non-linear

Networks as Ubiquitous

- Social network
  - who talks to whom
- Knowledge network
  - who knows what
  - (or a variation s.a. Resource network)
- Belief network
  - who believes what

Construct Evolves Meta-Networks

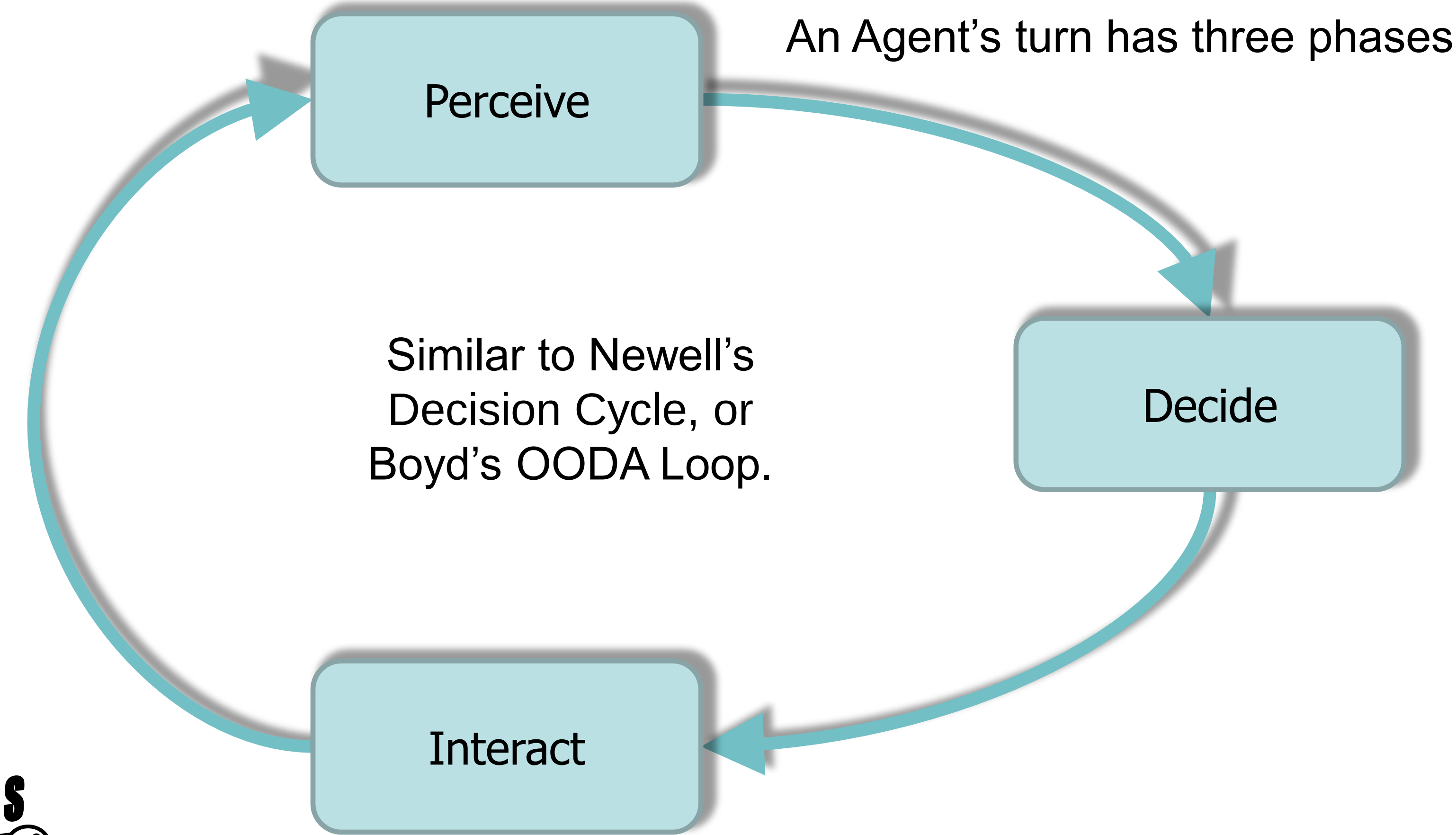
Meta-Network structure represented by
multiple networks of people, knowledge & tasks

	People	Knowledge	Tasks
People	**Social Network** Who knows who	**Knowledge Network** Who knows what	**Assignment Network** Who does what
Knowledge		**Information Network** What types of knowledge are related	**Needs Network** What knowledge is needed per task
Tasks			**Precedence Network** Which task must be done before which

(* adapted from Krackhardt & Carley, 1998)

Construct Agent Cycle

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



Bounded Rationality: Group schemas inform alter perceptions

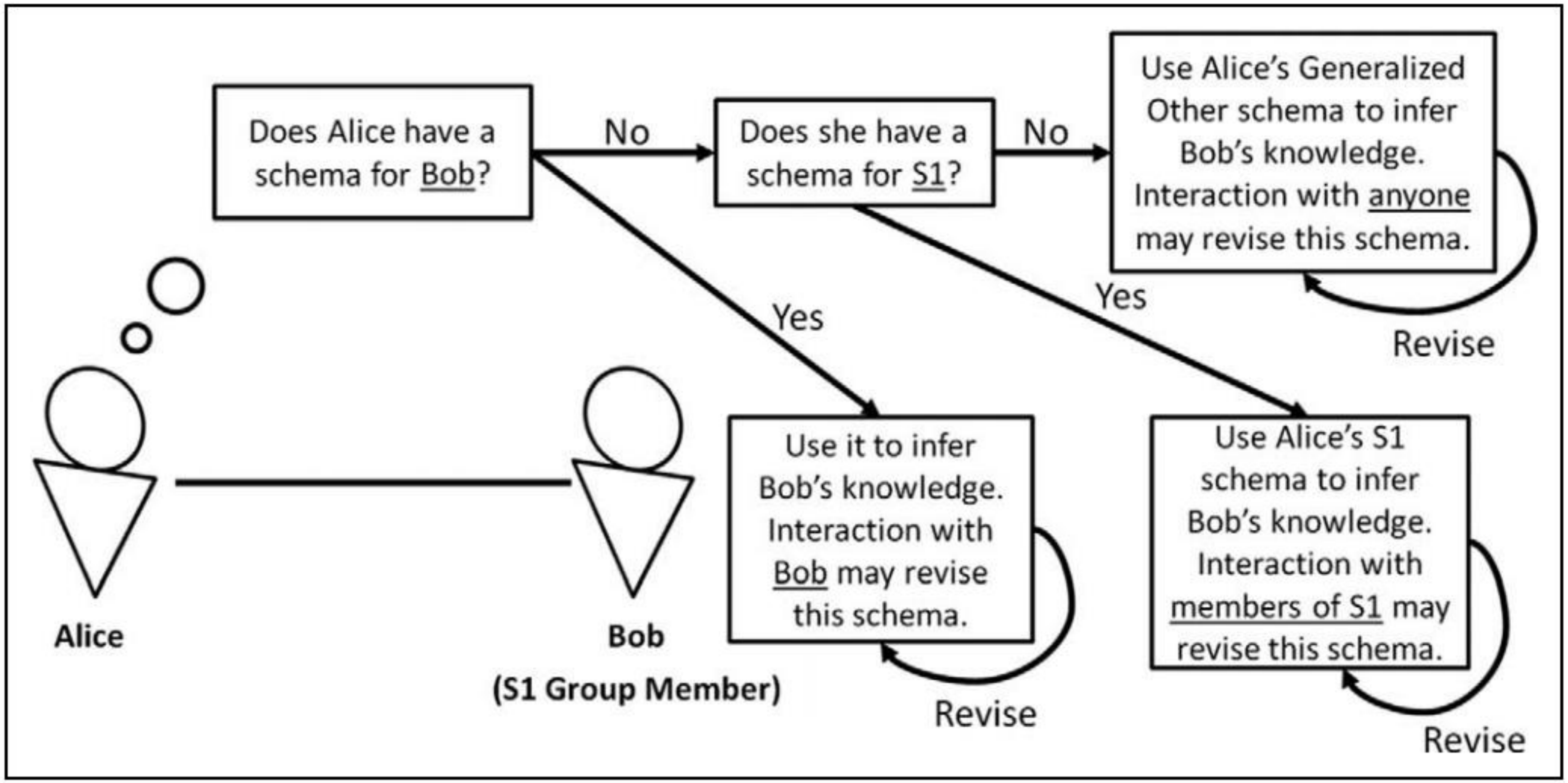


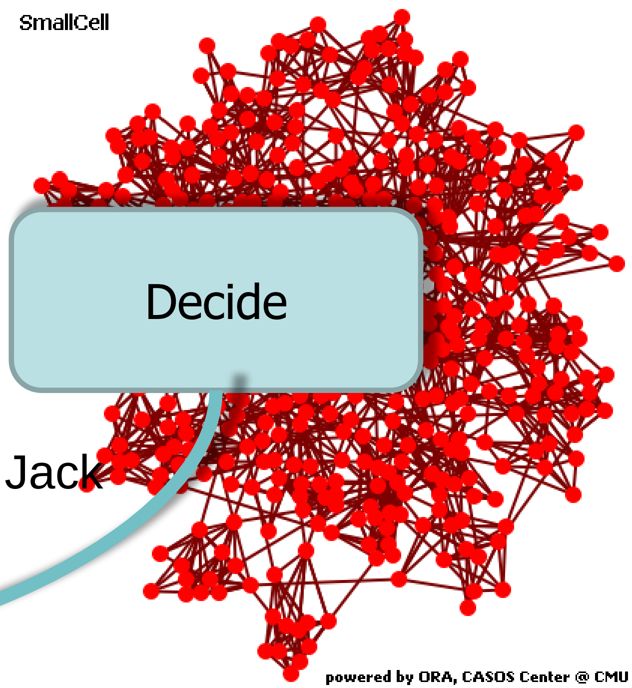
Figure 1. Alice needs to infer Bob's knowledge to determine her likelihood of interacting with him.

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.

Information Diffusion

- 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.



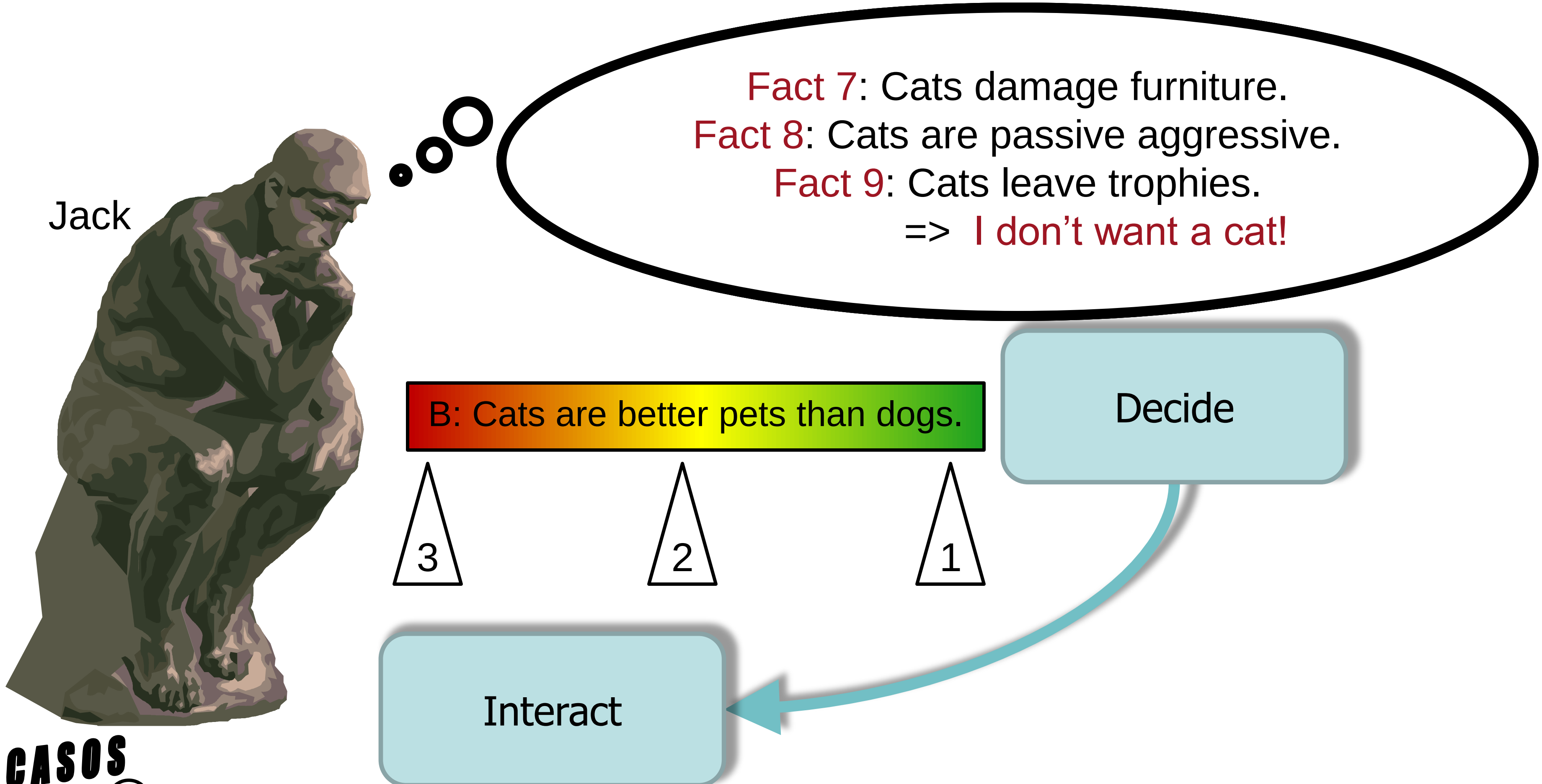
What the agent chooses to communicate can be:
- Knowledge ("Cats live longer than dogs.")
- Perception of another actor's knowledge
  ("Jack knows cats live longer than dogs.")
- Perception of belief
  ("Jack thinks cats are better pets than dogs.")
  The agent will attempt to infer what knowledge Jack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

Decide

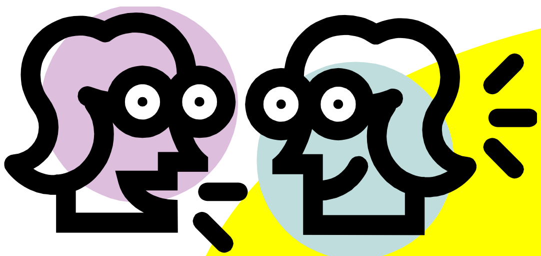
Interact

Belief Diffusion

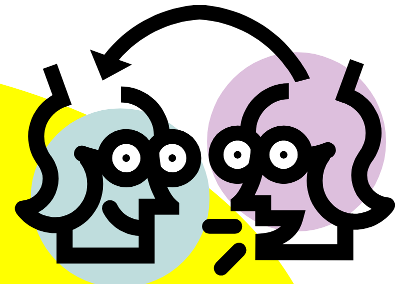
- Construct is used to model **belief diffusion**.



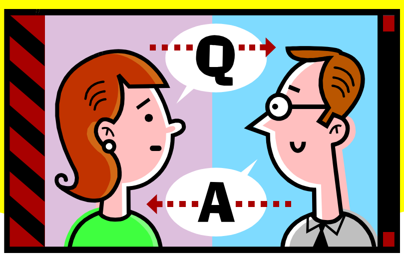
Agents Can Interact for Many Reasons



Homophily (Similarity)

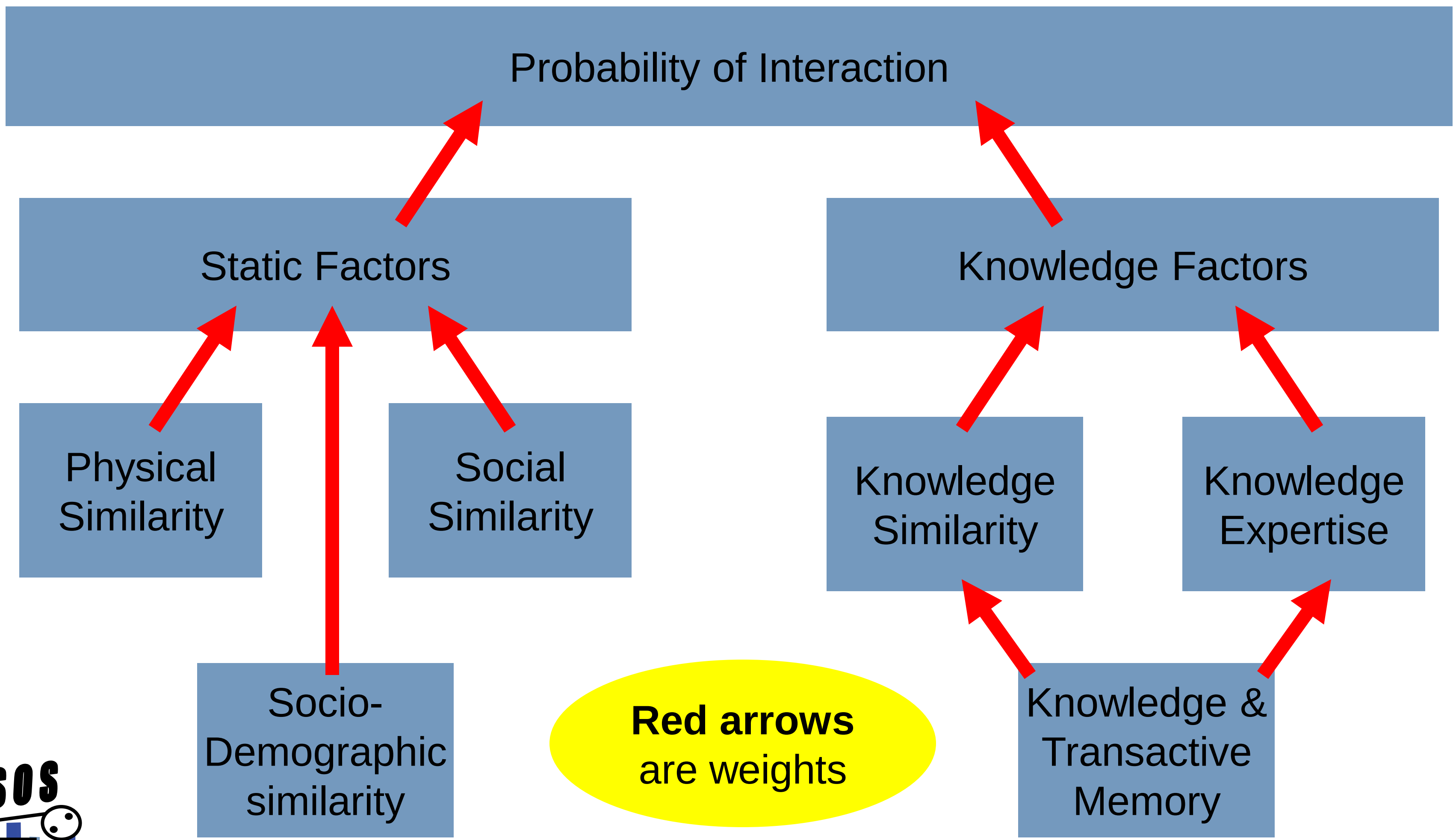


Learning New Knowledge

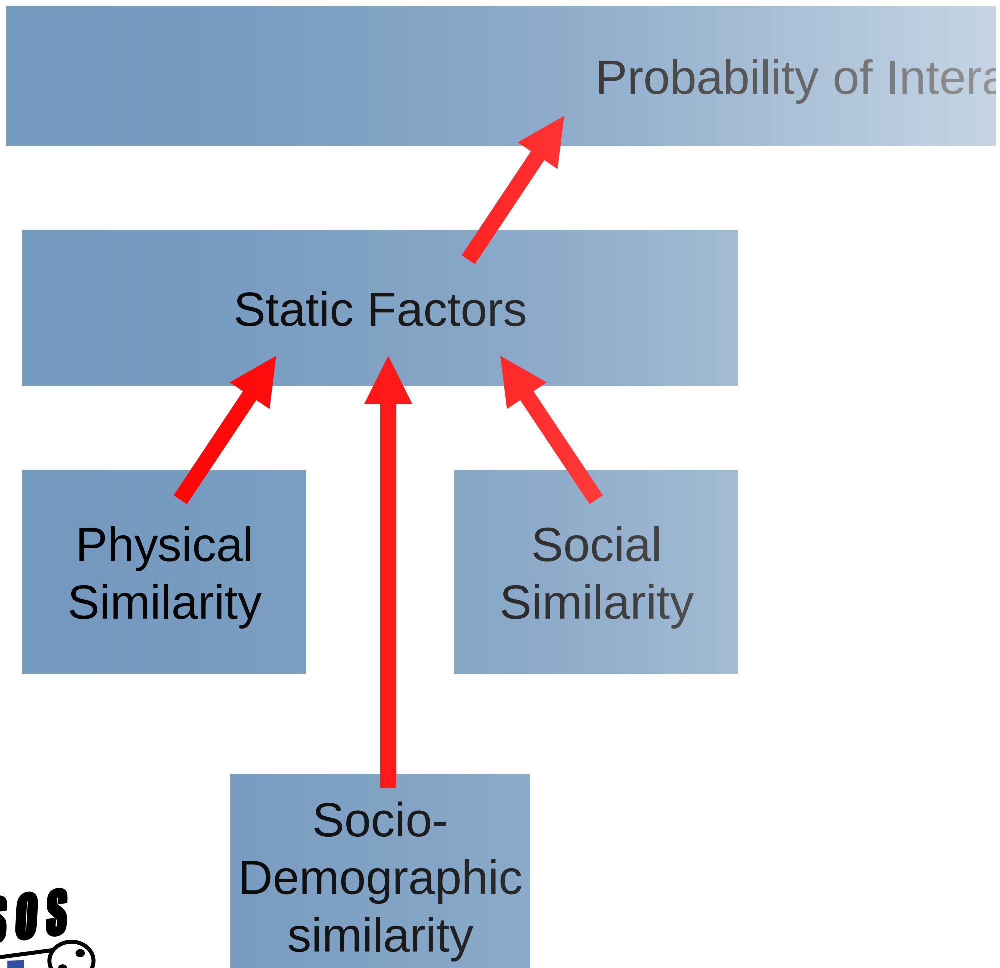


Explicit Information Search

Computing Interaction Probabilities



Static Factors

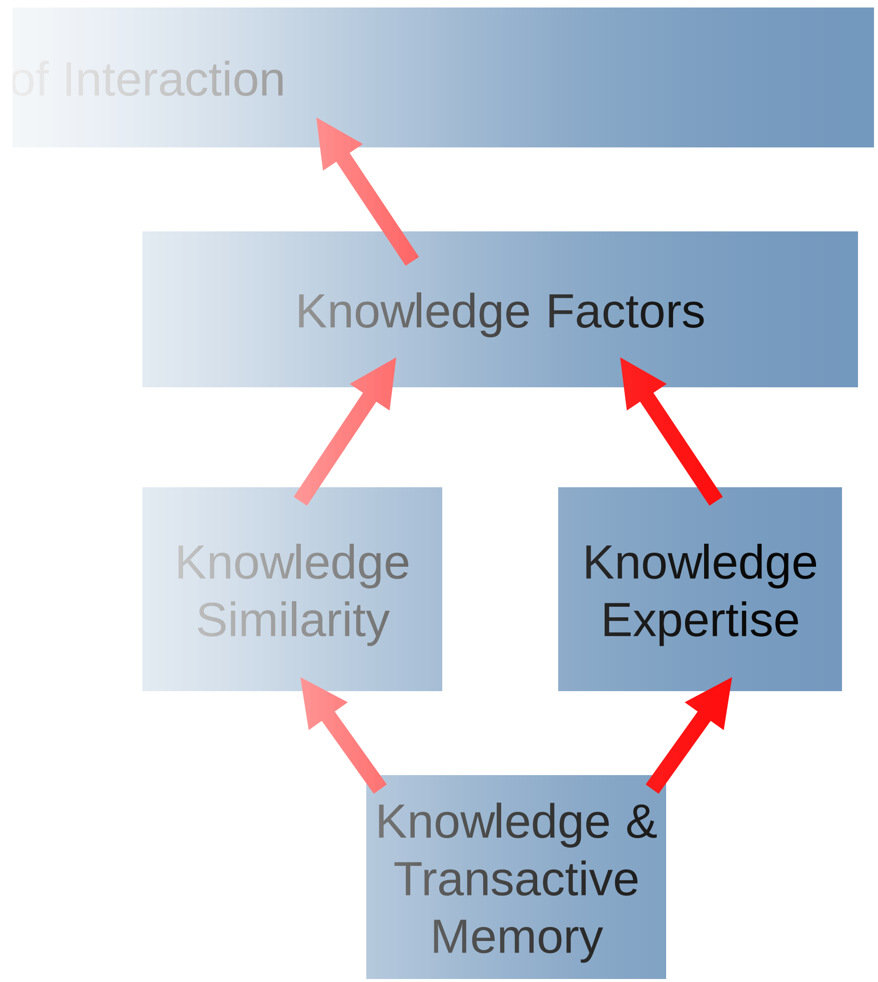


- Physical similarity
  - how close two agents are
  - can use a distance function
  - closer -> higher probability
- Socio-demographic similarity
  - computed from attributes
  - represent salient qualities
  - overlap -> higher probability
- Social similarity
  - catch-all for other features
  - can be ignored if not needed
  - more -> higher probability

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Knowledge Factors

- Knowledge similarity
  - how much info in common
  - more -> higher probability
- Knowledge expertise
  - how much other knows
  - more -> higher probability
- Knowledge weights
  - make some facts important
  - affect the calculation of simi-larity and expertise scores

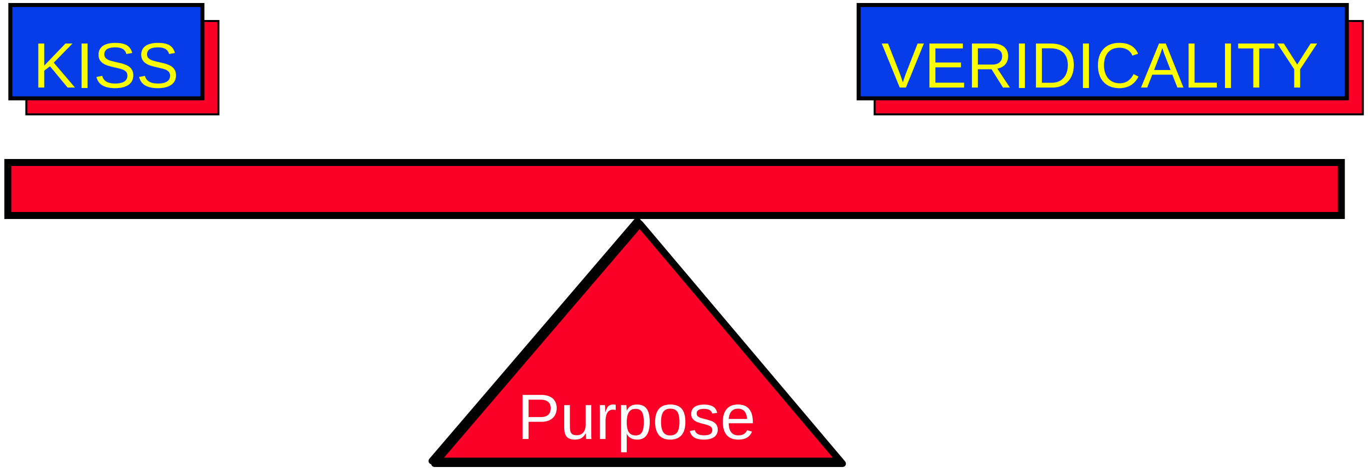


Where can I read more?

- Original Description
  - Kathleen Carley, 1991, "A Theory of Group Stability." *American Sociological Review*, 56(3): 331-354.
  - Kathleen Carley, 1990, "Group Stability: A Socio-Cognitive Approach." Pp. 1-44 in Lawler E., Markovsky B., Ridgeway C. & Walker H. (Eds.) *Advances in Group Processes: Theory and Research*. Vol. VII. Greenwhich, CN: JAI Press.
- Illustrative Uses and Validation
  - David S. Kaufer & Kathleen M. Carley, 1993, *Communication at a Distance: The Effect of Print on Socio-Cultural Organization and Change,* Hillsdale, NJ: Lawrence Erlbaum Associates.
  - Kathleen M. Carley & David Krackhardt, 1996, "Cognitive inconsistencies and non-symmetric friendship." *Social Networks*, 18: 1-27.
  - Craig Schreiber and Kathleen Carley, 2003, The Impact of Databases on Knowledge Transfer: Simulation Providing Theory, NAACSOS conference proceedings, Pittsburgh, PA.
  - Craig Schreiber and Kathleen Carley, 2003, Going Beyond the Data: Empirical Validation Leading to Grounded Theory, NAACSOS conference proceedings, Pittsburgh, PA. First runner up for the NAACSOS graduate student paper award.

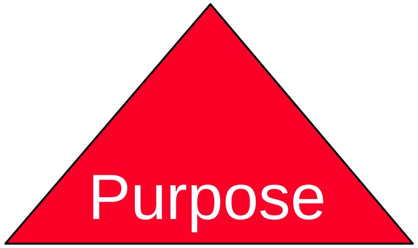
MODEL DEVELOPMENT

Balance



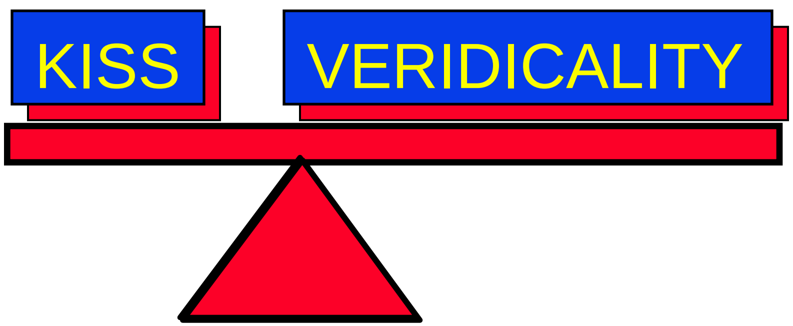
IMPORTANT!!!!!

Different types of models require different evaluation schemes & levels of validation

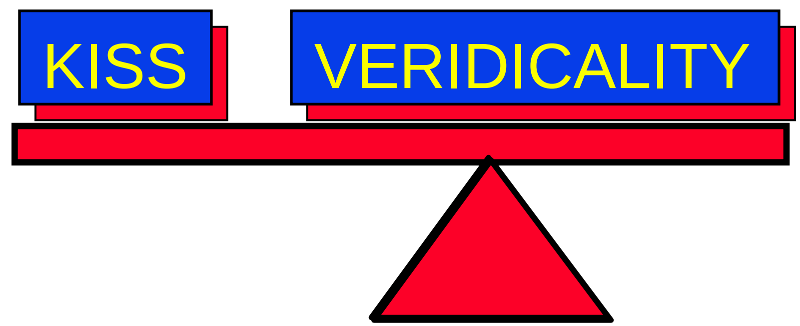


Shifting the Balance

The purpose of the model may shift the balance between simplicity and veridicality.



Intellective or theory building models



Engineering, emulation, or wind tunnel application models

Kitchen Sink Approach

People buy into model and suggest changes

Ease of adding features

Lack of a minimal theory

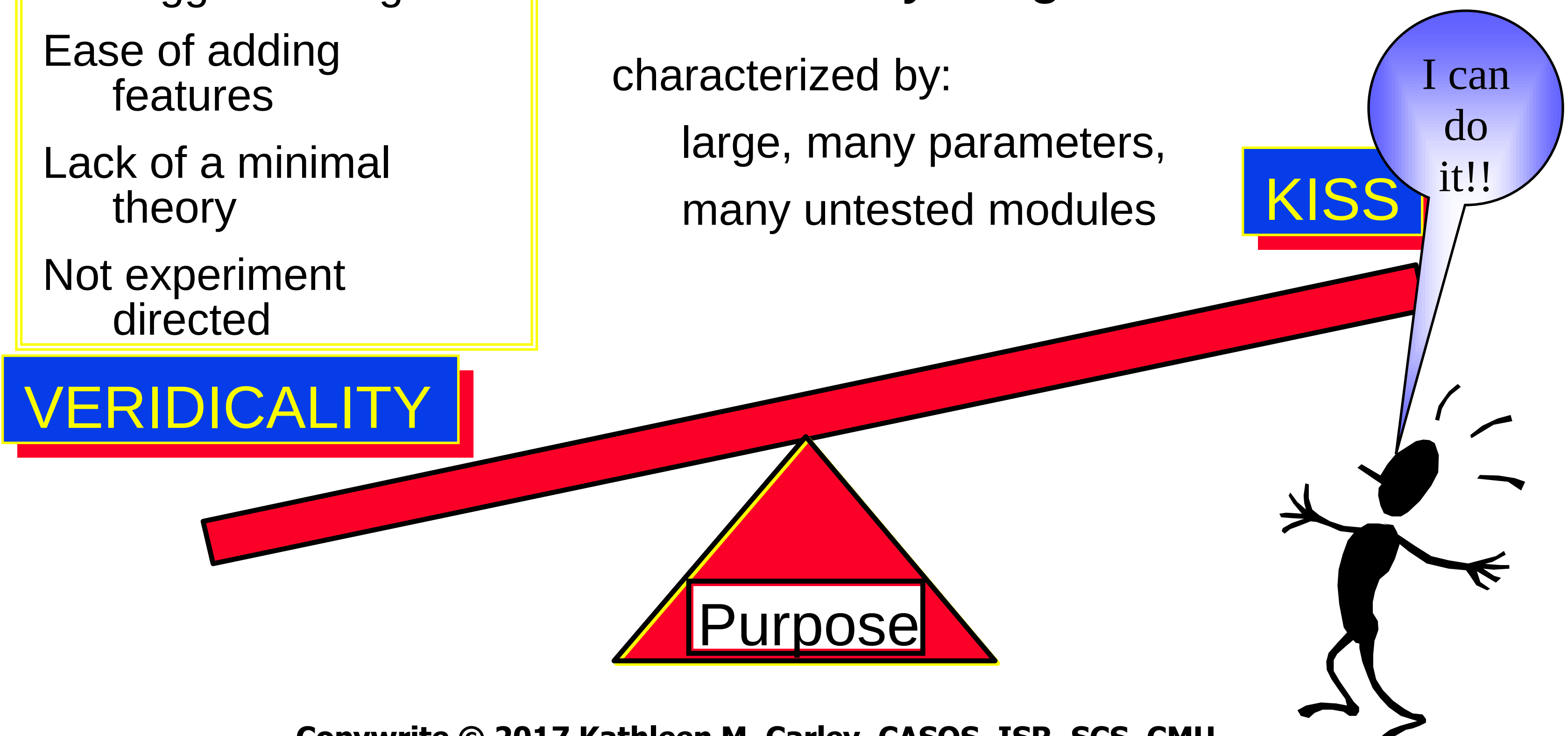
Not experiment directed

Everything in the model and model everything

characterized by:

large, many parameters,

many untested modules



Illustrative Approach

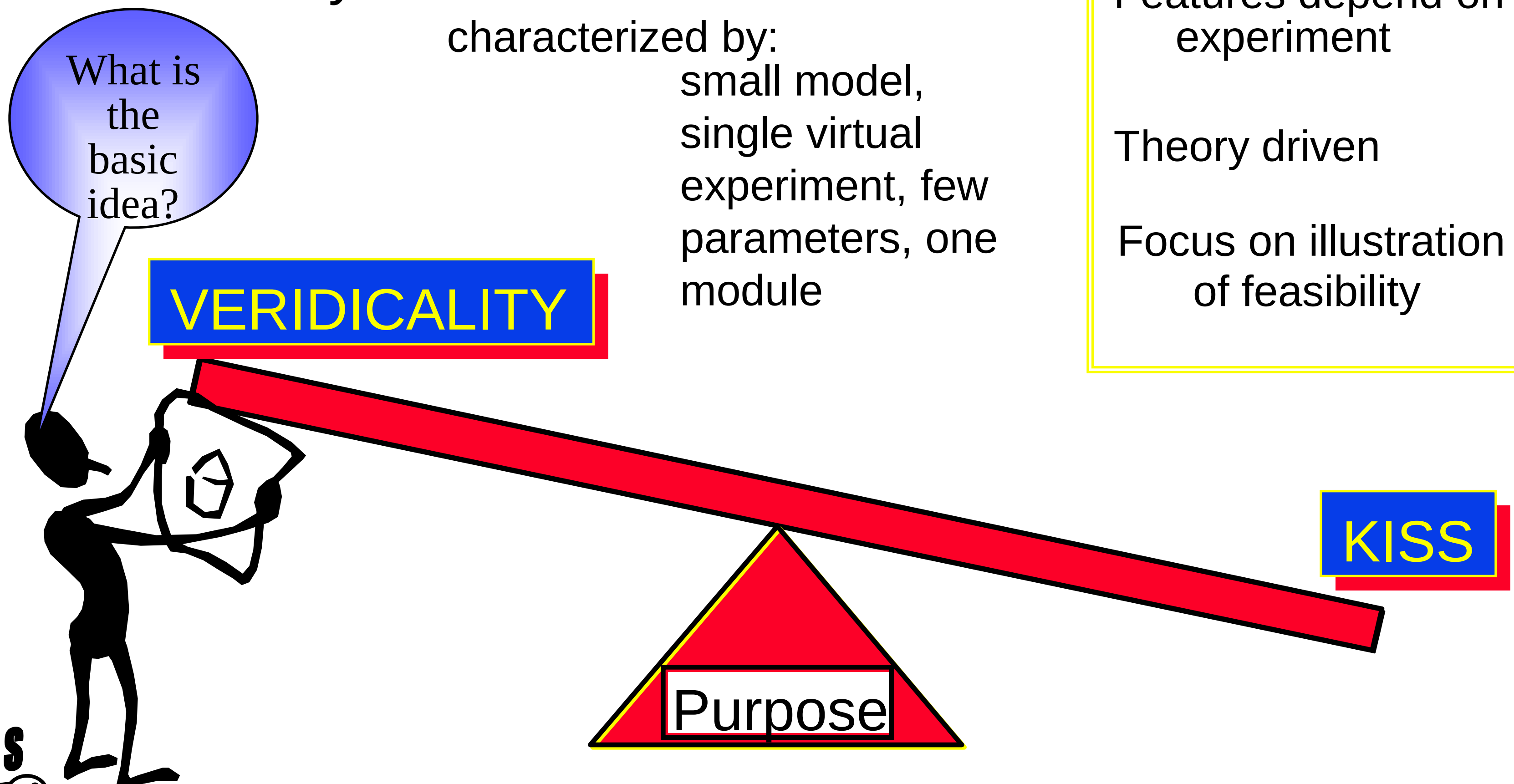
Parsimony and illustration rule

characterized by:
small model, single virtual experiment, few parameters, one module

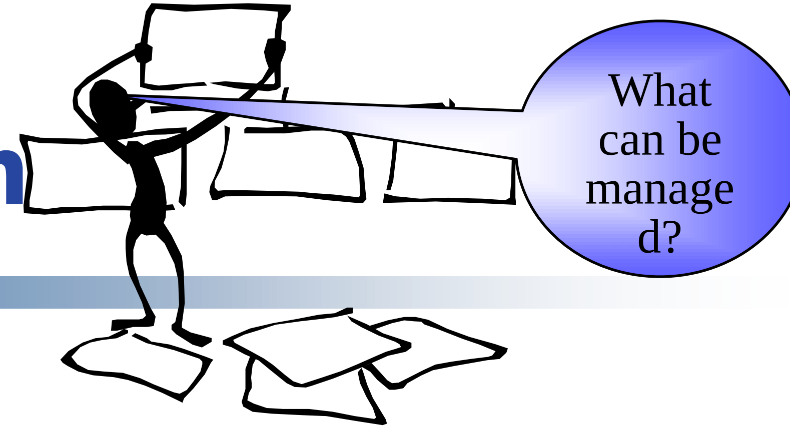
Features depend on experiment

Theory driven

Focus on illustration of feasibility



Emulative Approach

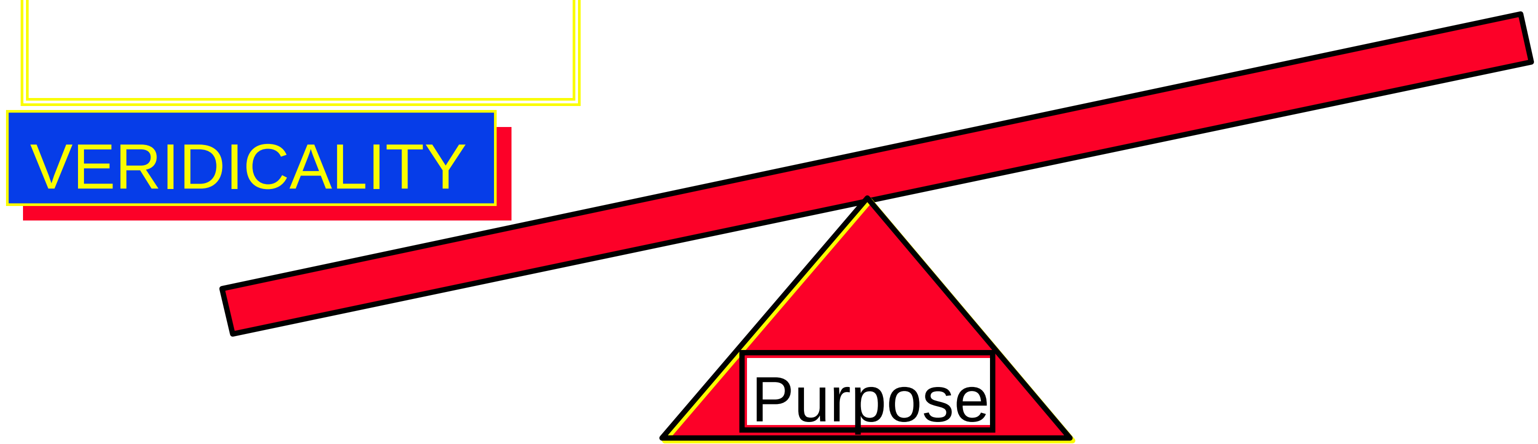


Built via blocks
Empirically grounded
Addresses applied issues

Sufficient unto the task

characterized by:

large, partly tested modules



Taking a Building Block Approach

Build & Test

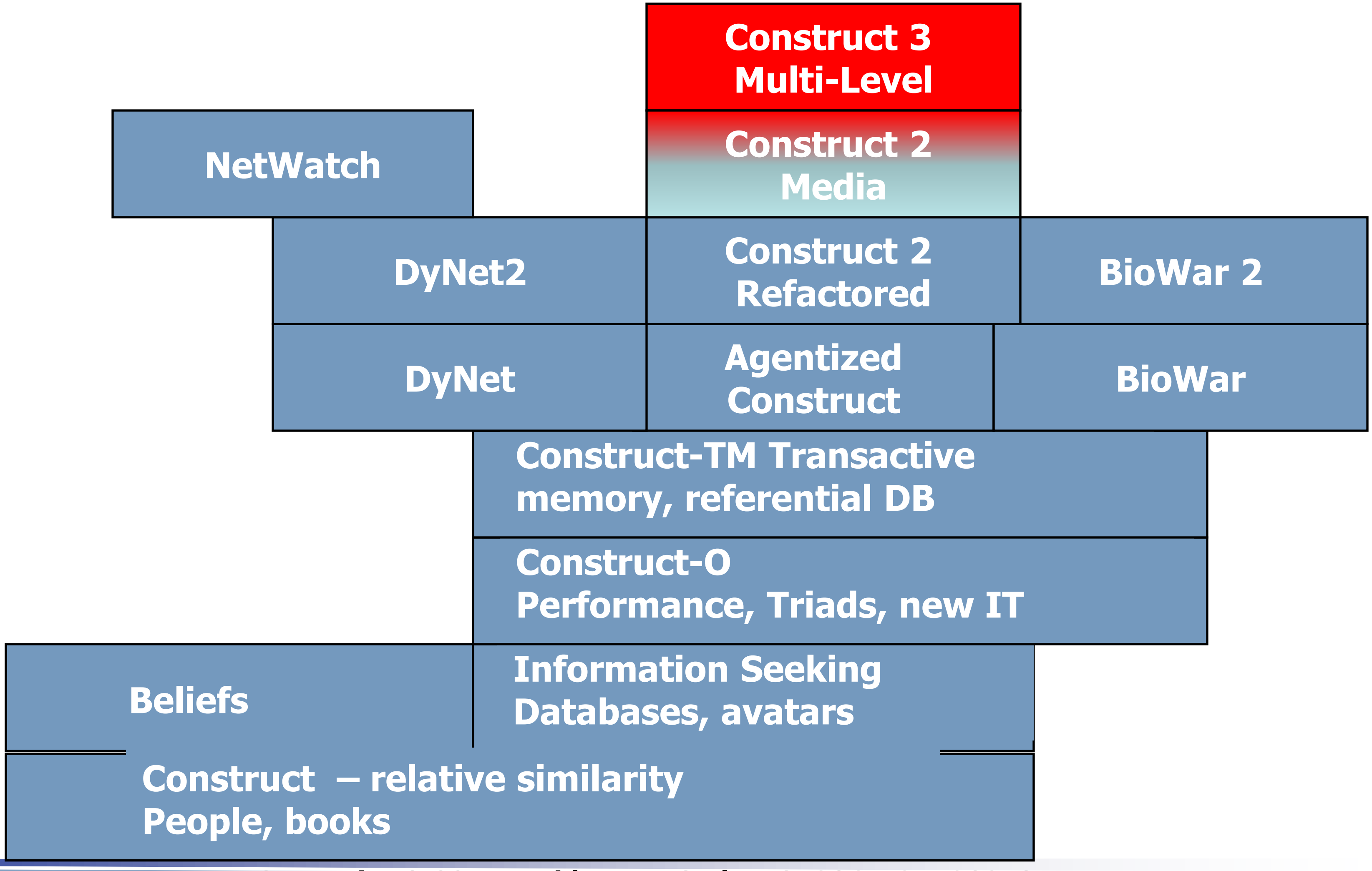
Build & Test

Build & Test | Build & Test

- Start out small
- Keep only those parts necessary for the theory
- If it is a complex theory focus only on the core
- If a needed component is not in the theory treat it as a parameter don't invent mechanisms
- If you don't know how something changes begin by assuming that it does not
- Build, run a virtual experiment
- Add critical variable

Note: this approach is facilitated by using modular programming and object oriented programming techniques

CONSTRUCT: Building Blocks



Data-Centric Modeling

James Hardy/AltoPress/Maxppp

Improve veridicality while still keeping model mechanics simple

META- AND MULTI-MODELING

What is Multi and Meta Modeling?

- Meta-Modeling is the idea that multiple models exist for similar purposes, and that meta-modeling analysis can be used to provide the "best-fitting" model given data and goals.

- Multi-Modeling is a method to combine various pre-existing models, often from different modeling frameworks, to address a complex problem.
  - Different Models, but similar IO: Validation through triangulation
  - Different Models, and different IO: Extended conclusions

- While simple in concept, the "devil is in the details"

Meta-Modeling

- Meta-Modeling requires
  - A stable of models
  - Criteria by which there is a best model for a given purpose

- Example Criteria
  - Prior validation
  - Intended Use
  - Units
  - Scale

Why not just big complex models?

?

Issues with (Social) "Models of Everything"

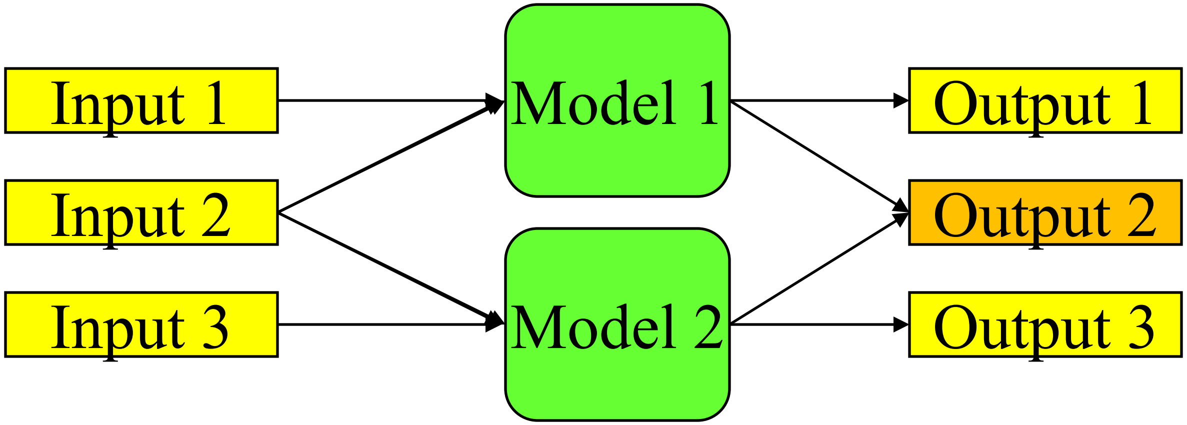
- Model maintenance
- Complexity of interactions
- Clarity of results
- Input and Output choices/complexity
- Model Use Problems
- Model Interpretation Issues

MULTI-MODELING: CONFIGURATIONS

Multi-Modeling - Linking Models Together

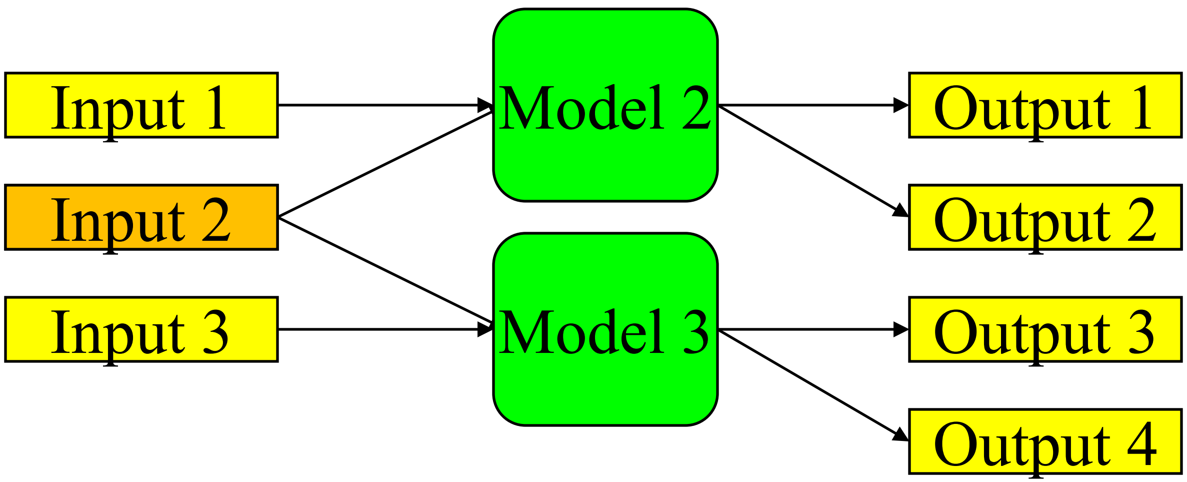
- Given two models, we have several ways of lining them up with each other based on their inputs and outputs
  - Docking
  - Collaboration
  - Interoperation
  - Integration

Docking



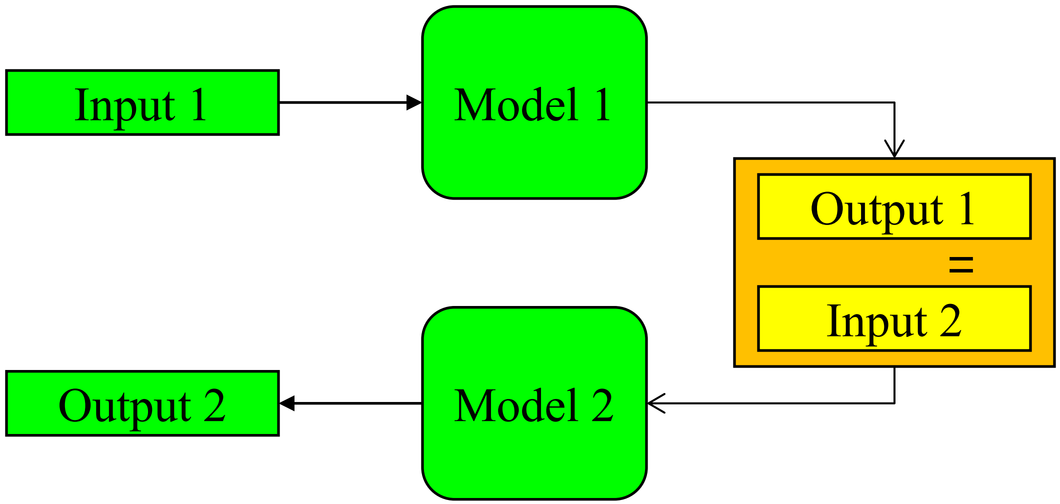
- Pro:
  - Economical
  - Flexible
  - Extensible
  - Form of validation
- Con:
  - Requires initial developers or great documentation
  - Low theory development
  - Low reuse

Collaboration



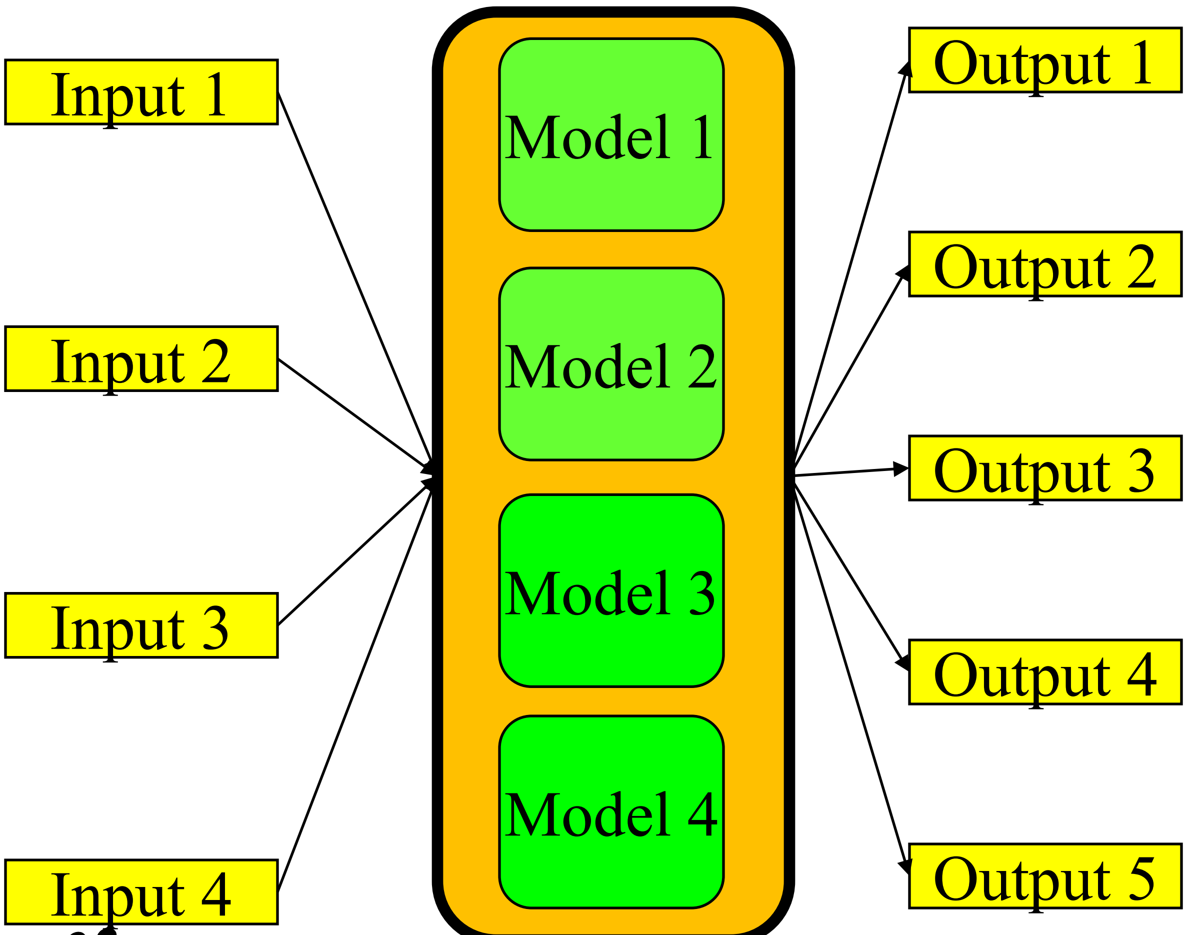
- Pro:
  - Economical
  - Flexible
  - Extensible
  - Form of validation
  - Supports use of non-fused data
- Con:
  - Moderate Theory Development
  - Low reuse

Interoperation



- Pro:
  - Easy to Use
  - Extensible
  - Flexible
  - High reuse
  - Form of verification & validation
  - Economical
  - Some Theory development
- Con:
  - Some Code Development
  - Some Theory development

Integration



- Pro:
  - Easy to Use
- Con:
  - Costly
  - Code Development
  - Theory Development
  - Increased difficulty to validate
  - Low extensibility
  - Low flexibility
  - Low reuse

Configuration Summary

Configuration	Inputs	Outputs
Docking (A|B)	Aligned	Aligned
Collaboration (A+B)	Aligned	Not Aligned
Interoperation (A>B)	Model Output	Not Aligned
Integration ([AB])	Integrated	Integrated

MULTI-MODELING ISSUES (WHY IS THIS HARD?)

Model Access

- Can you get the model?
- Example “good” inputs and outputs of the model?
- Example “bad’ inputs that cause model failure?
- Can you get source-code?
- Do you have access (and then attention) of the model developers?

I/O Alignment

- Question 1: Construct Validity
  - Are we talking about the same "stuff"?
  - Example: Is Model A's "wealth" output the same as Model's B "prosperity index" input or are they, at least, highly correlated?
- Question 2: Units
  - "I didn't know we were in the metric system."
  - "Oh, in my graphs, I always show this number as Millions of Units, but the model reports the actual number."
- Question 3: Parsers and Post-Processors
  - "My input parser assumes the data is comma-delimited, because I know there will be no commas in any of the input, ever."
  - "Oh, my post-processor transforms that number to this more understandable number, but I'm sending you the first number."

Error Propagation

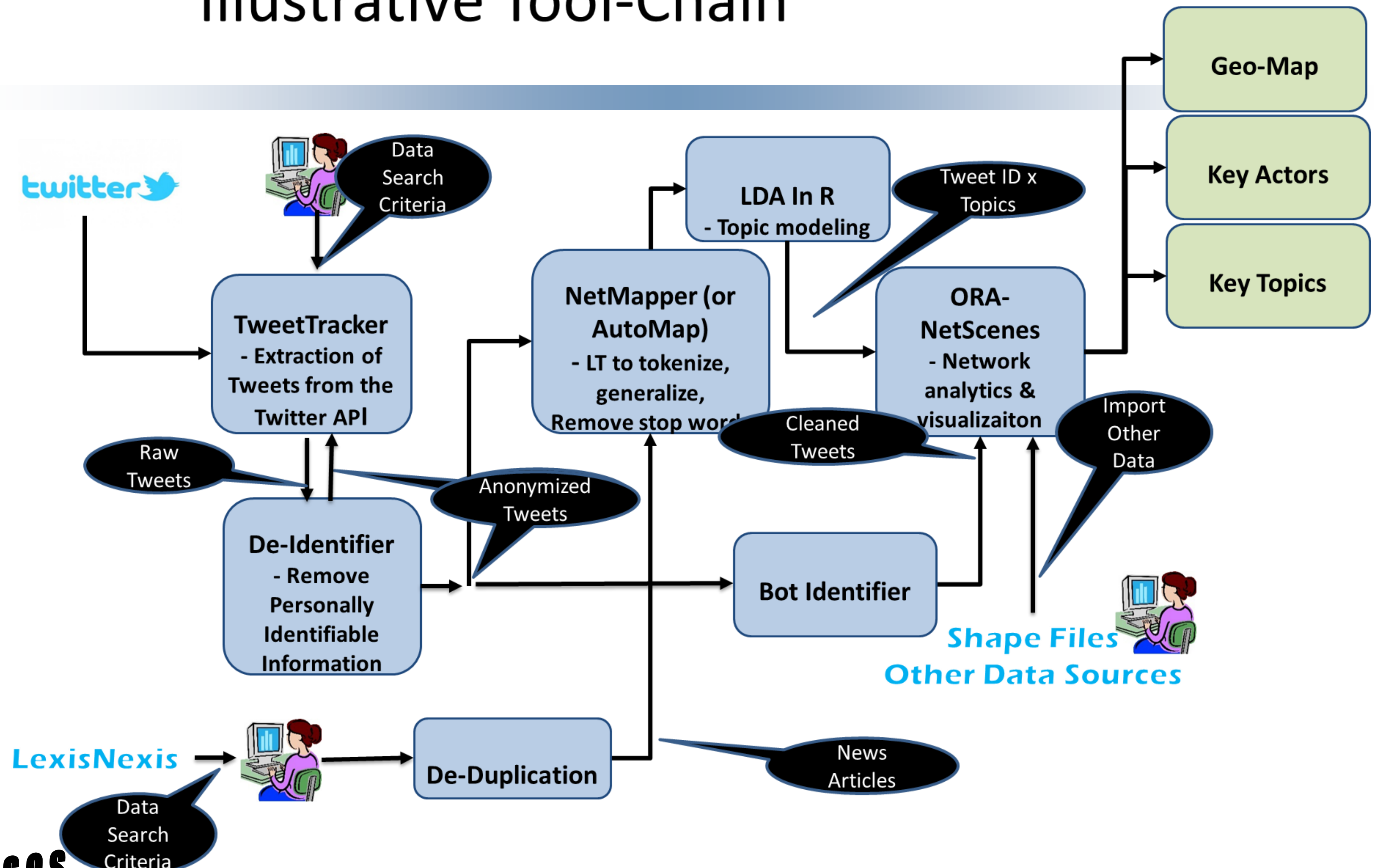
- Very difficult to nail down the sources of failure when multiple models are linked together.
- Example:
  - Model A has a rare edge condition where it will produce questionable output, the output looks valid, but the outcomes are wrong. This edge condition wasn't included in the original validation exercise of Model A.
  - Model B is linked to Model A through inter-operation, and thus relies on Model A's output. The inter-operated model includes states where the rare edge condition will occur, and Model B will report bad results due to Model A's output. Since Model B is "last" in the result, most developers will focus initial debugging on B.
- This is another good reason to do "validation in parts"

Scale & Boundary Conditions

- Temporal
  - *Are the models operating in the same time window?*
    - Yes – docking or interoperation or integration
    - No – collaboration or interoperation
- Group
  - *Are the models taking into account the same actors?*
    - Yes – docking or interoperation or integration
    - No – collaboration or interoperation
- Geo-Temporal
  - *Are the models considering the same spatial region in the same time frame?*
    - Yes – docking or interoperation or integration
    - No – collaboration or interoperation

MULTI-MODELING: EXAMPLES

Illustrative Tool-Chain



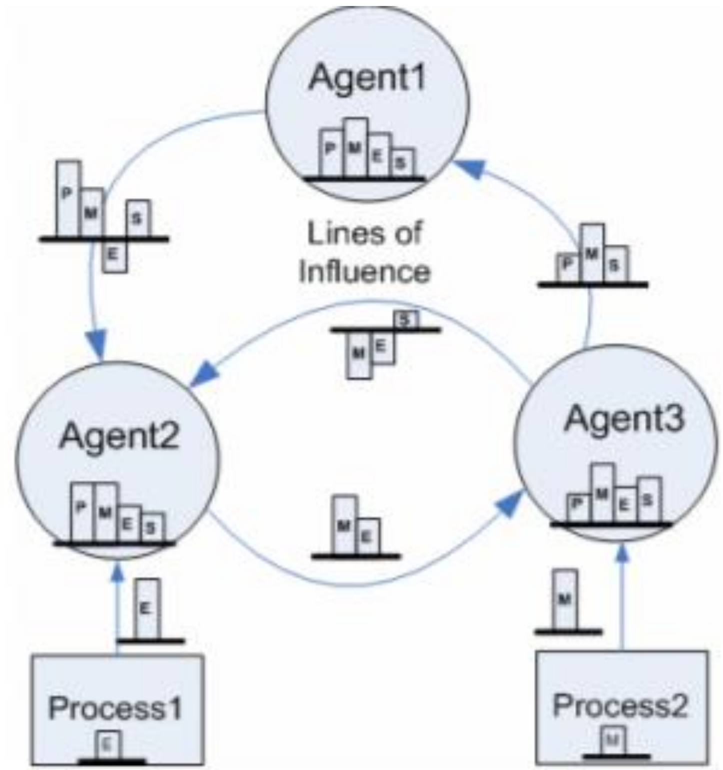
PSTK – Power Structure Tool-Kit

- The PSTK (Power Structure Tool-Kit) was a model designed for military planners to use to understand secondary impacts of FSO (full-spectrum operations)

- These were its goals/requirements
  - Must be usable by non-programmers (A “No PhD” rule)
  - Must be based on good social, political, and cognitive theory
  - Must be easily applied to new problem areas
  - Must be able to model groups at different granularities

Model Operation

Agents choose how to spend power along available "lines of influence" based on goals. In effect, a hybrid system-dynamics agent-based model.



Processes are links to other models and can be sources or sinks.

Integration with Other Models

- PSTK Models operated in conjunction with many other models via a shared data backplane

- Note the MRM itself was turn-based and resolved all interactions between models before models **could proceed to their next turn**

Sorta like this, but with more headaches

C2 Wind Tunnel

- The C2 Wind Tunnel was developed as a means of evaluating human-in-the-loop command and control technologies.

- Multi-University Project:
  - Vanderbilt
  - George Mason
  - CMU
  - UC Berkeley
  - Naval Postgraduate School

Integration through GME

- All tools to be added to the C2 Wind Tunnel must work with the Generic Modeling Environment (GME) specification
- The GME is an example meta-modeling, where a sufficiently precise ontological description of the model, including:

"*The modeling paradigm contains all the syntactic, semantic, and presentation information regarding the domain; which concepts will be used to construct models, what relationships may exist among those concepts, how the concepts may be organized and viewed by the modeler, and rules governing the construction of models. The modeling paradigm defines the family of models that can be created using the resultant modeling environment.*"

- From "Command and Control Wind Tunnel: Integration and Overview" by Karen E. Roth and Shelby K. Barret

Carnegie Mellon
isr institute for SOFTWARE RESEARCH

C2 Overview

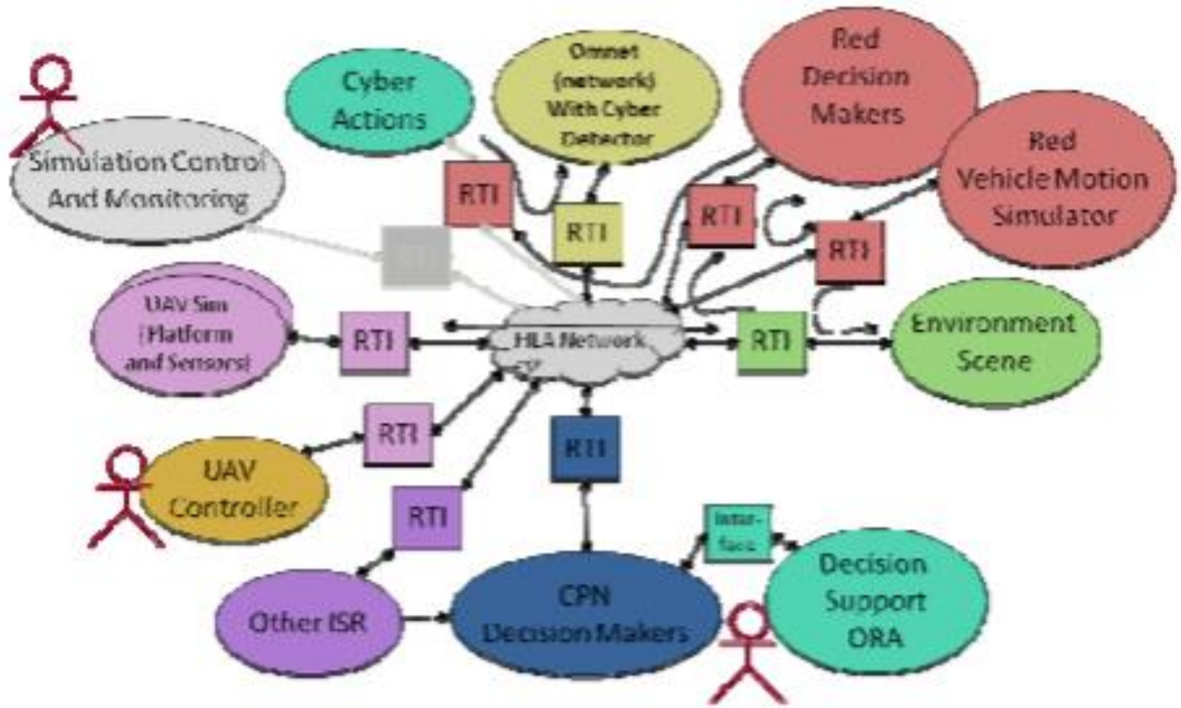
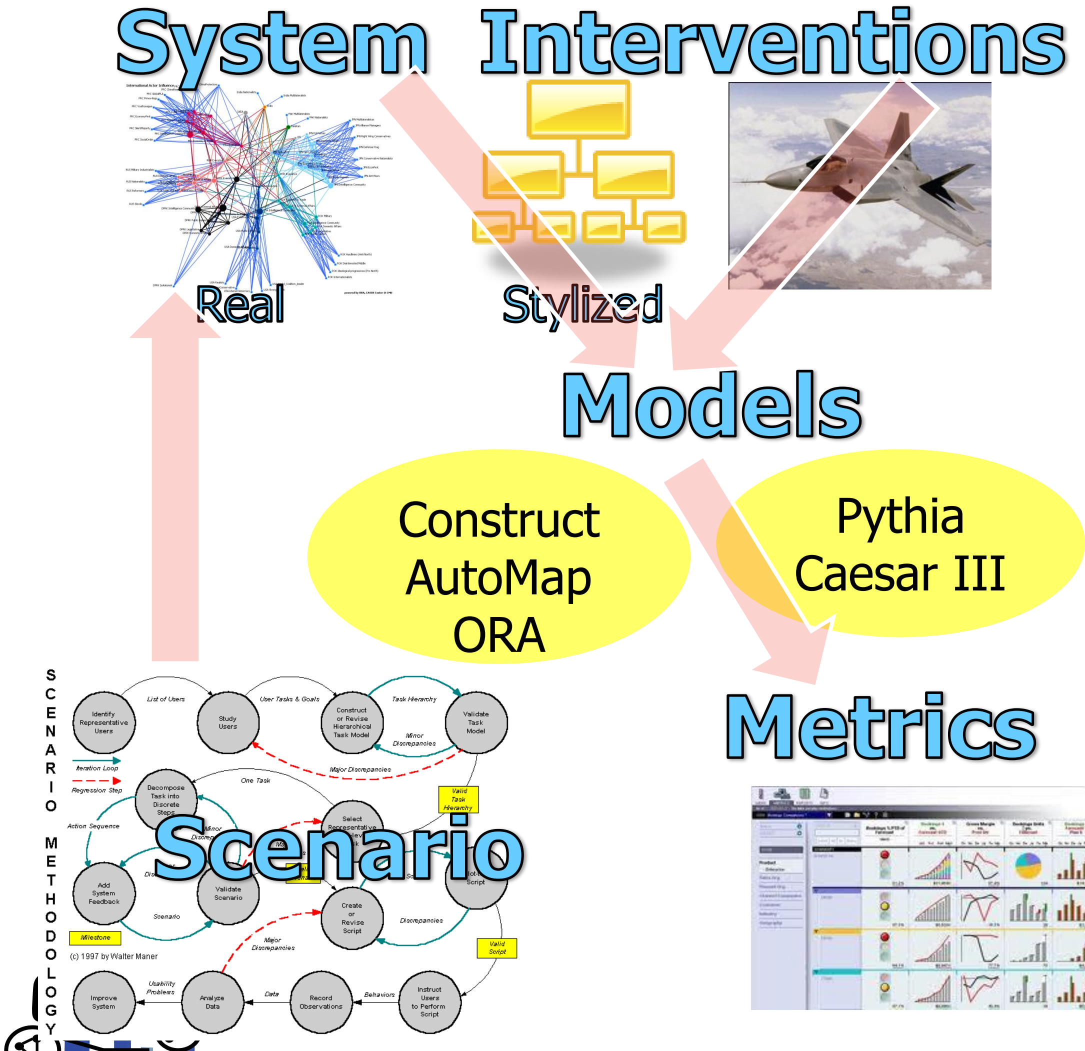


Figure 3: High Level View of C2 Wind Tunnel Components for Demonstration [6]

CASOS

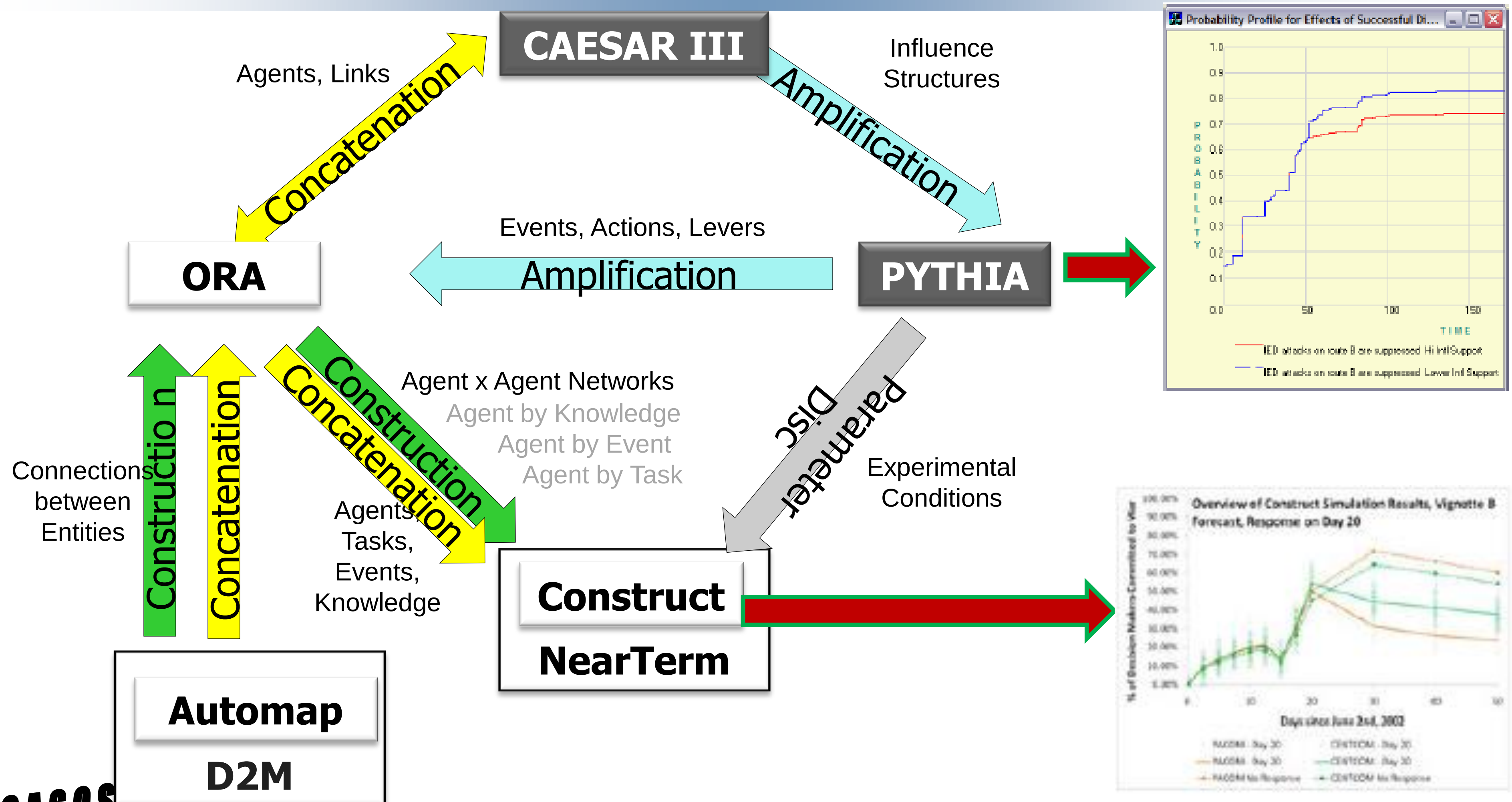
Socio-Technical System Analysis of Nuclear Deterrence



Approach & Methods

- Discovery Science
  - Induction
  - Hypothesis Generation
  - Translational Research Techniques
- Scenario Direction
- Data Driven Models
- Model Driven Experimentation
- Multi-Modeling

Nuclear Deterrence Example

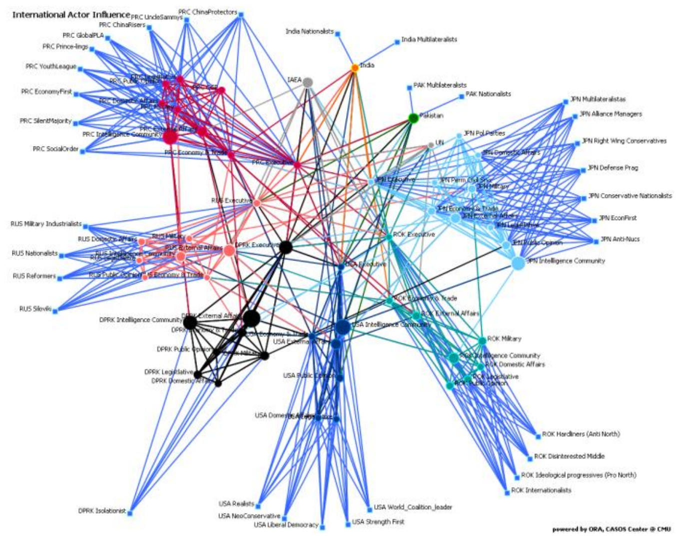


Role of Networks: Nuclear Deterrence

SME Data

ORA

Network Analysis



Construct

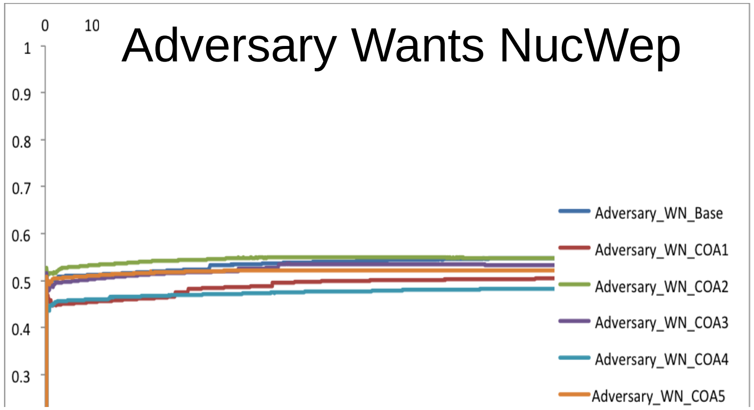
ABM

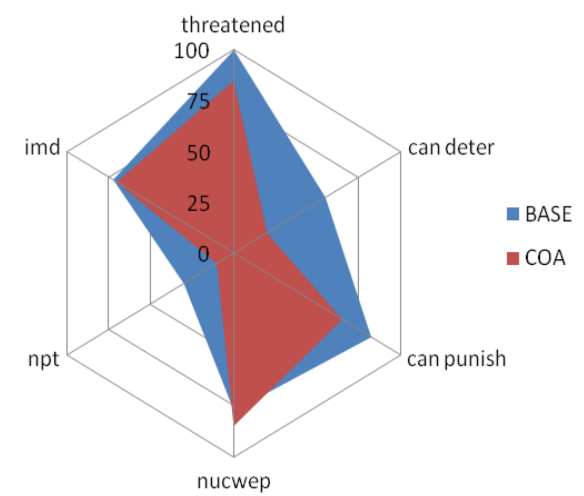
PYTHIA

TIN

Actor	Rank	Betweenness
DPRK External Affairs	1	724.68 (0.25)
USA Intelligence Community	2	643.37 (0.22)
JPN Intelligence Community	3	525.67 (0.18)
DPRK Executive	4	461.00 (0.16)
DPRK Intelligence Community	5	444.23 (0.16)

Illustrative Model Results





- Generic Model:
  - None of the COAs dramatically improved or degraded agents' nuclear weapons posture.
  - Example – all COAs led to slight increase in adversary's desire for nuclear weapons
  - Consistent with Pythia's outputs.
- Regional Model:
  - COAs change DPRK's nuclear posture
  - Example - doubling IMD capabilities of US, ROK, and JPN increased DPRK's desire for nuclear weapons
- Models support further experiments
  - Fine and Course Grained actions by US, Allies, Adversaries, and other parties
  - Varying quantities of stakeholders, special agents, knowledge, core beliefs and beliefs of interest

MULTI-MODELING: SUMMARY

Multi-Modeling lets us work together

- Multi-Modeling is a great way to bring together a joint research team at multiple institutions to tackle an interesting problem.
  - Each research group is responsible for their models and for providing guidance for good I/O
  - Since you're solving a hard/complex problem, lots of opportunity for novel research and publication

- Funders love joint projects

Multi-Modeling can be a nightmare

- Issues of boundary spaces, construct validity, and I/O issues can halt progress on any multi-modeling effort
- It's important that models, or at least toy models with similar I/O, be provided to other research groups as early as possible
- Never underestimate the amount of integration time required

It's Models – All the Way Down

VIRTUAL EXPERIMENTS

What are Virtual Experiments?

- Virtual Experiments are like experiments, but they involve testing a model rather than testing reality.
  - We keep the descriptor "virtual" to remind ourselves that these are not 'truths' we are discovering, but instead truth as the model predicts.
  - The assumptions of the model must always be kept in mind when considering the results of the model.
  - "All models are wrong, but some are useful."

- Many of the same principles involved in designing an experiment must also be kept in mind when designing a virtual experiment.

When to use Virtual Experiments?

- When testing with real people is:
  - Too expensive
  - Unethical
  - Infeasible

- You should not use it:
  - When you can get what you want from a survey (most surveys are going to be cheaper than building and testing a model)
  - When you're looking for 'truth' and not 'trends'

Caveats to VEs

- There are serious, sometimes skirted issues with computational models. You’re not working with real people, so
  - You have to buy the model of how people work to buy the results
  - You have to code the model to get the results
  - Writing code without bugs isn’t easy

Model Assumptions Exercise

- Exercise: Given what you know from the previous discussion on Construct, collaborate with one or more nearby people and try to come up with a list of assumptions for Construct.

- Assumption (from Logic): an **assumption** is a proposition that is taken for granted.

- Also, if something can be a parameter (or an input to the model) then it's not a model assumption (although it may be an experimenter assumption)

(Some) Key Assumptions in Construct

- Agent-Orientation – The phenomenon of interest can be modeled through the actions and interactions of individual agents (of whatever grain size).
- Network Representation – Connections between actors can be represented as one or more networks.
- Homophily – Agents prefer to interact with other agents similar to themselves.
- Expertise-Seeking – Agents prefer to interact with agents that have rare knowledge.
- Turn-Taking – Agents are relatively equivalent in their ability to react to events, and thus fixed time blocks are appropriate.

Model Assumptions Exercise 2

- Exercise: For each of the following assumptions, collaborate with one or more others nearby and try to identify phenomena where that assumption **does not** hold.
  - Agent-Orientation
  - Network Representation
  - Homophily
  - Expertise-Seeking
  - Turn-Taking

Virtual Experimentation

- As stated before, many of the same problems of general experimental design come up in Virtual Experimentation

- Dependent Variables
- Independent Variables
- Method (non varying but still need to be set parameters)
- Control Conditions
- Generality
- Power (repetitions)

Independent Variables

- What am I changing?
- For Virtual Experiments, this should both be the variable name, and what values you intend it to have.
- **Be careful of combinatoric explosion – too many independent variables and it'll take 100 years to run your simulations.**

Dependent Variables

- What am I measuring?
- Is what I am measuring a good analog to the thing I want to measure (in the real world)?
- Do I have some reason to believe that what I'm manipulating will change the values of what I'm measuring?

    But it's not a direct manipulation!

- **Its easy to think of tens of metrics you want to think about. There is likely one or two that best fit your RQ

Method

- For Virtual Experiments, much of the ‘method’ is in setting variables that are not being manipulated but still must be specified.

- There are three reasonable strategies for these variables
  – Set them so they don’t have any impact
  – Set them to a reasonable base-line
  – Have the variable set randomly across an appropriate distribution

When would you use any one of these methods?

Control Conditions

- Not really the same as in a standard experimental design.
- In a Virtual Experiment – control conditions are settings of the independent variables least likely to have any effect on the phenomena of interest.
- With network topologies, ER Random networks are often used as a control condition for topologies.
  - This is despite the fact that ER Random networks are not very realistic!

Generality

- Defining model parameters can become very specific – the source of parameters should always be drawn from literature.
- Example: Examining network information flow after actor removal.
  - Bad example:
    - Case 1: Remove Gordon
    - Case 2: Remove Jill
    - Case 3: Remove Pat
  - Good Example:
    - Case 1: Remove Actor with highest degree centrality
    - Case 2: Remove Actor with highest betweenness centrality
    - Case 3: Remove Actor with highest eigenvector centrality

Power

- Given enough repetitions, even trivial differences between simulation conditions will produce statistically significant results.
- It's important to focus on trends, rather than specific values.
  - Wrong: Because of the manipulation condition, Y increases by 5%.
  - Better: Y tends to increase under the manipulation condition.
- A reasonable heuristic is 25 repetitions per combinatoric

Example

How does varying the degree of ethnocentrism in an artificial society affects the formation of social relationships across social groups under different models of the underlying cultural structure?

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*.

Virtual Experiment

Parameters	Values Taken
Parameters of Interest	
Initial knowledge distribution	random, group based, all same
Initial Bias Parameter (IBP)	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1
Other Parameters Varied	
Group Activation Threshold (GAT)	-5, -1
Group Learning parameter (GLP)	5, 25, 50
Individual Activation Threshold (IAT)	-1, 0
Constants	
Number of Simulation Turns	150
Number of Agents	1000
Number of Knowledge Bits	500
Number of Interactions	2
Number of Knowledge bits passed per interaction	1
Density of knowledge (percent of bits set to 1)	0.4
Decategorization Parameter (DP)	6
Groups Per Agent	1
Total number of groups	4
Repetitions	
Number of repetitions	10
Total Runs	3*11*2*3*2*10 = 3960

Analyzing the results

- Run the simulation
- Construct a network of who talked to who more than N (N=2 here) times
- Look at the *log-odds* of a tie to a member of the outgroup

$$log_2\left(\frac{\#relations\ connecting\ two\ agents\ in\ different\ groups + 1}{\#relations\ connecting\ two\ agents\ in\ the\ same\ group + 1}\right)$$

Results from VE

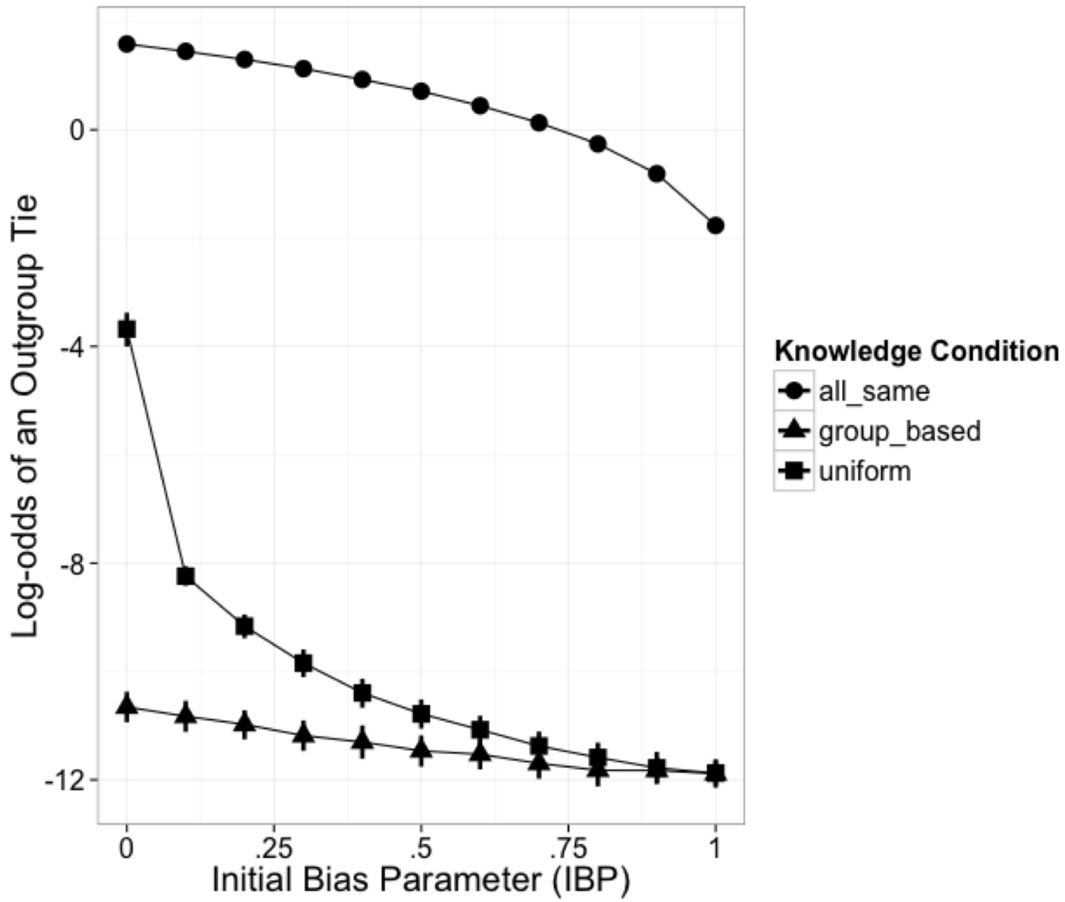


Figure 4- The x-axis represents the ten different IBP conditions and the three different shapes of points represent knowledge conditions. The y-axis gives the log-odds of an out-group tie, and lines connect the mean outcomes across the different conditions. Ninety-five percent (95%) bootstrapped confidence intervals are drawn at each IBP condition.

Results from VE

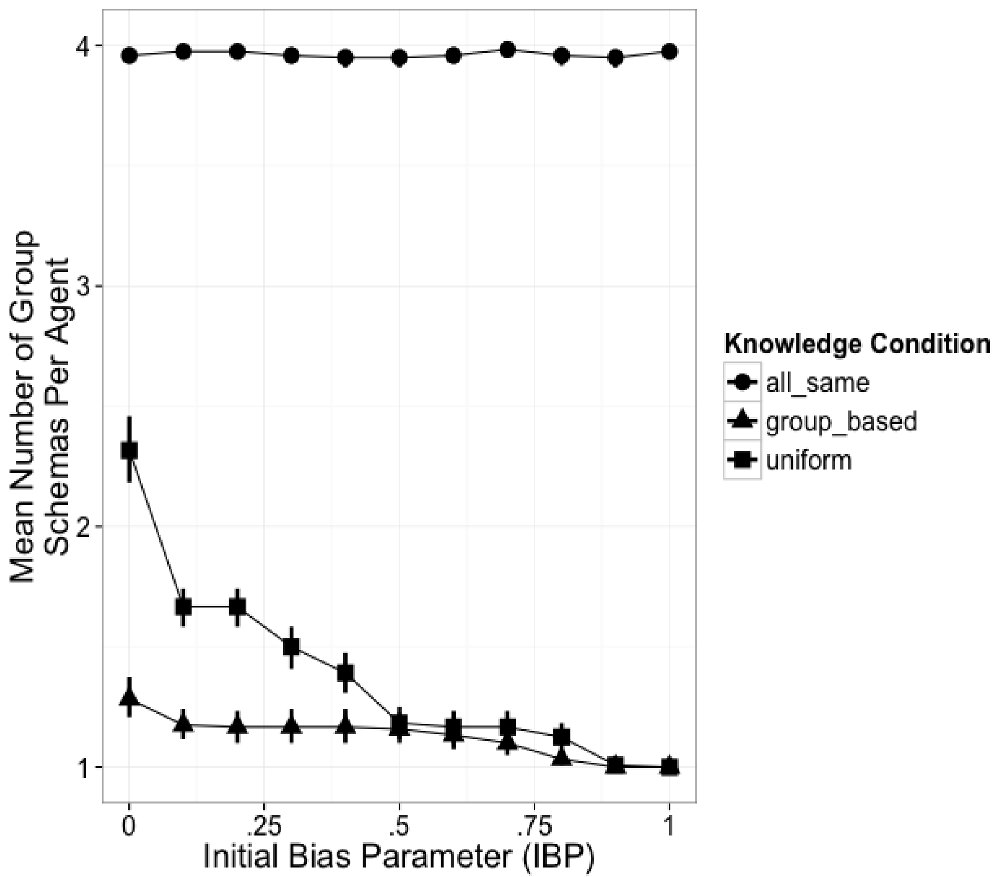


Figure 5a) The mean number of group schemas that agents held across all conditions. Error bars are 95% bootstrapped confidence intervals

Conclusions from VE

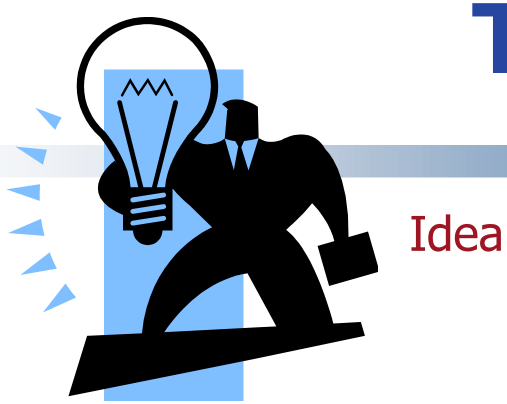
- The model generated results that were sort of obvious
  - In my eyes, this is a good thing!
  - What do you think?
- Results suggested that neither stereotypes nor the form of underlying cultural structures alone are sufficient to explain the extent of social relationships across social groups
- Rather, we provide evidence that shared culture, social relations and group stereotypes all intermingle to produce macro-social structure.
- What do you think should be next?
  - Cross-cutting groups
  - Differentiating in-group love from out-group hate

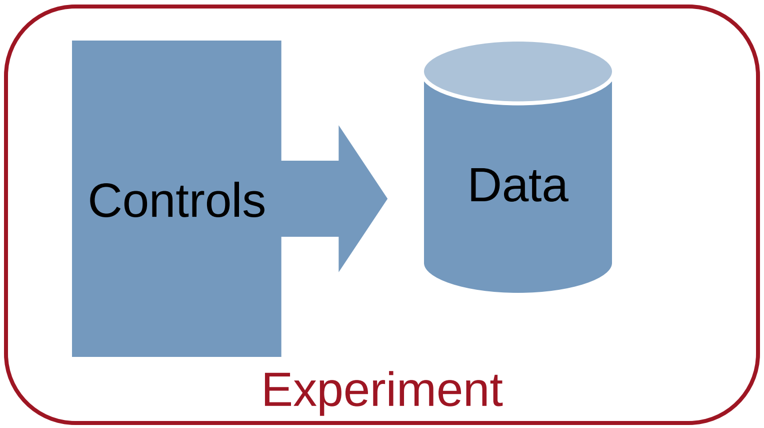
VALIDATION

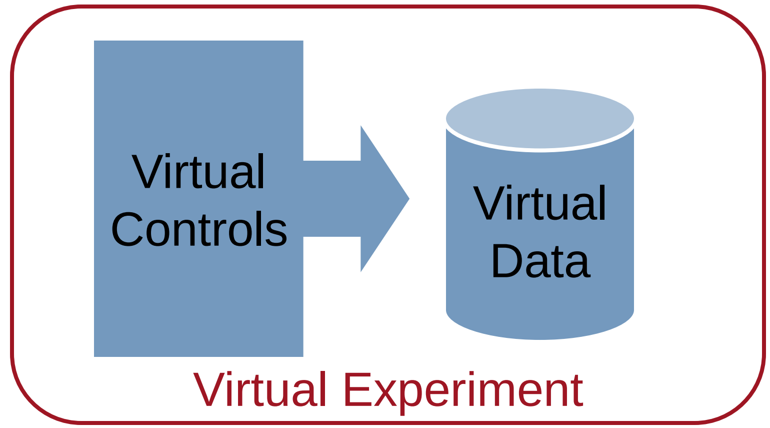
Key Questions

- Should Every Model be Validated?
- How much validation is needed?
- What is the difference between validation and curve fitting?
- What is validation?

The Theory







Assumptions:
Experiment is repeatable
Results can be replicated
Data is complete
Controls are stable and well defined

Assumptions:
Code is accurate
All factors have been modeled
Standard statistical tests can be used

VV&A

- Verification
  - Is the programming correct
  - *Did I build the thing right?*
  - *The process of determining that a model implementation and its associated data accurately represent the developer's conceptual description and specifications*
- Validation
  - Does the model match reality at a level sufficient for the model’s purpose
  - *Did I build the right thing?*
  - *The process of determining the degree to which a model and its associated data provide an accurate representation of the real world from the perspective of the intended uses of the model*
- Accreditation
  - What systems can this model be used on and in what context
  - *Should it be used?*
  - *The official certification that a model, simulation, or federation of models and simulations and its associated data is acceptable for use for a specific purpose*

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

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.

Validation Levels

- 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

Simple 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.

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

Multi-expert Problem

- What if experts or cases disagree?
- Typical solution: average the two cases
- Alternative: put in both cases as options with a certain probability of occurring
  - Probability:
  - Equally weighted
  - Weight can reflect degree of agreement

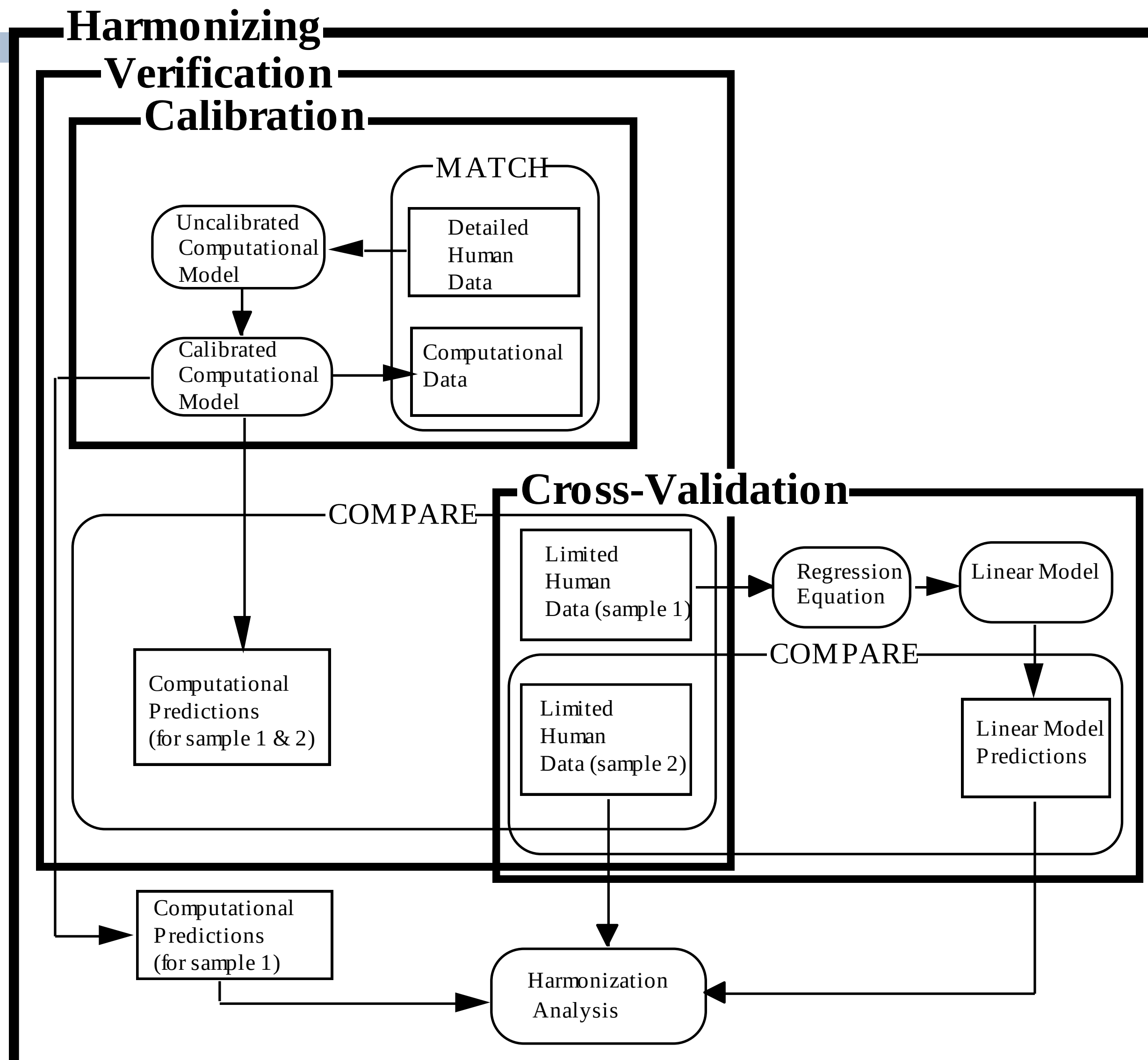
NOTE: For rule based models detailed cases may be the opinions of experts

Harmonization

- **Assessment of theoretical adequacy by comparison with a cross-validated linear model**
- **Harmonization involves contrasting the predictions of a computational model and a linear model**
  - **Requires enough cases that you have two samples large enough for statistical analysis**
  - **Requires that there is a reasonable linear model**
- **Harmonization can locate areas of the model where the embodied theory is inadequate**

statistical comparison process

Harmonizing



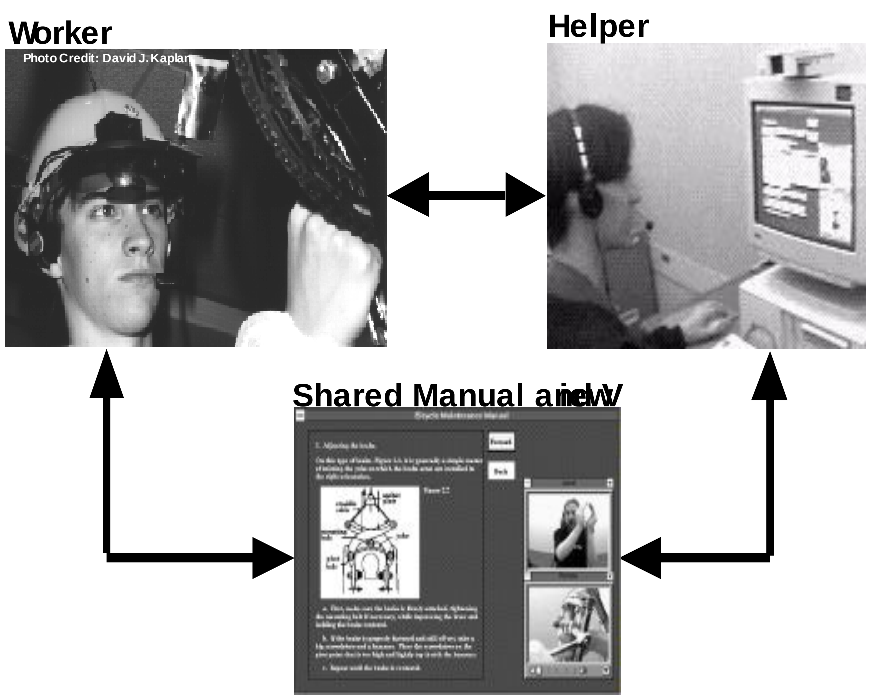
Locating the Linear Model

- Many sources for such a model:
  - Linear model of inputs
  - Easily collected data that might be used by management to make the same prediction the computational model is making
  - Model presented in the literature

Illustration & Harmonization

COMIT - impact of technology on task performance
Human data on a bicycle repair task using various technology (including vide an non-video)

Focus is on prediction frequency and order of communication actions

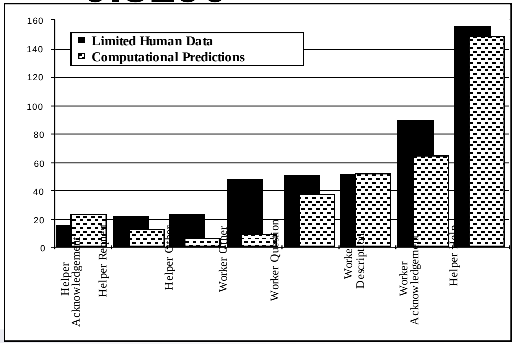


- Coded audio soundtracks
- Took half of subjects and created linear model and verified COMIT
- Linear model based on data accessible by managers
- Generated predictions for remaining subjects for COMIT and linear model
- Calculated R2 across subjects
- Data - 4 people, COMIT 2 people

COMIT Results

Action	Linear Model	COMIT	Correlation
Helper Acknowledgment	0.230	0.653	0.8083
Helper Request	0.498	0.141	0.7488
Helper Other	0.118	0.323	-0.3306
Worker Other	0.307	0.368	0.6886
Worker Question	0.215	0.011	0.4442
Worker Description	0.086	0.182	0.8148
Worker Acknowledgment	0.044	0.473	0.2986
Helper Help	0.310	0.130	0.8190

positive correlation means both models doing well/poorly in same areas, negative correlation means models tend to have opposite predictions



VALIDATION: BIOWAR EXAMPLE

BioWar

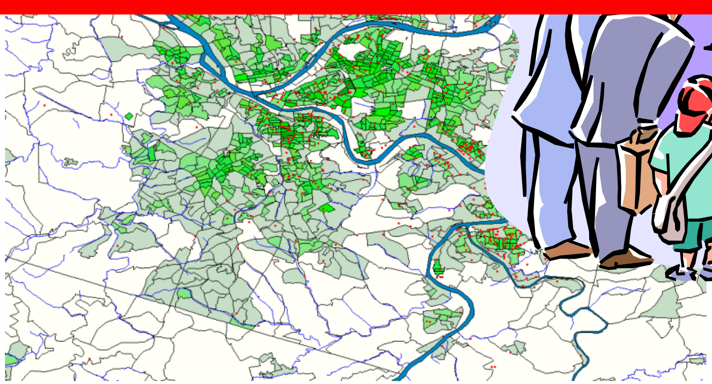
- City-level Multi-Agent Dynamic-Social-Network model of population response to loss-of-life events – bio-warfare, epidemiological, chemical
- High fidelity
- US Centric
- 5 cities/ 6 cities
  - Pittsburgh, San Diego, San Francisco, Washington DC, Norfolk/Hampton Roads
- 62 diseases
  - Weaponized – Smallpox, anthrax,
  - Background – influenza
  - Epidemic – pandemic influenza, SARS
- Sub-Model for military bases for use in planning, training and assessment of force vulnerability
- GenCity: Network generation: Census based imputation tool for realistic network generation that is context sensitive

The Model …

- Input
  - Military Bases
  - Census data – social, economic, occupation
  - School district data
  - Worksite and entertainment locat
  - Hospitals and clinics locations &
  - Social Network characteristics
  - IT communication procedures
  - Wind characteristics
  - Spatial layout
  - Disease models (symptom level)
  - OTC and Prescription drug info
  - Attack or event
  - Interventions
- Illustrative Output
  - OTC & Prescription drug sales
  - Insurance claim reports (Dr. visi
  - Emergency room reports
  - Absenteeism (school and work)
  - Infected, Contagious, Mortality

Agents move in networks which influence what they do, where, with whom, and what they know, what diseases they get, when, how they respond m, etc. Major difference in netwo l disease nd age.

Vast quantities of data
Real and Virtual
Format of virtual to match real
But real data is …
Incomplete
Diverse sources
Inconsistent
Different levels of granularity



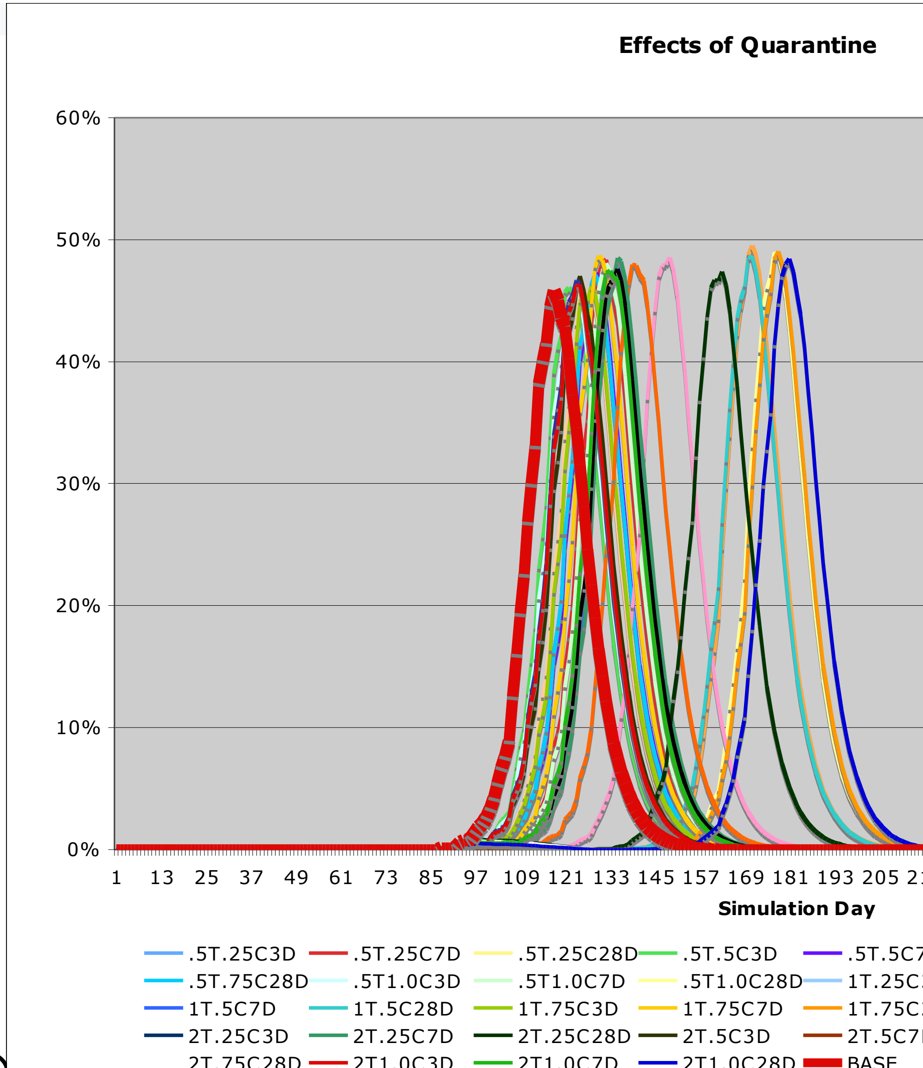
Model Objectives

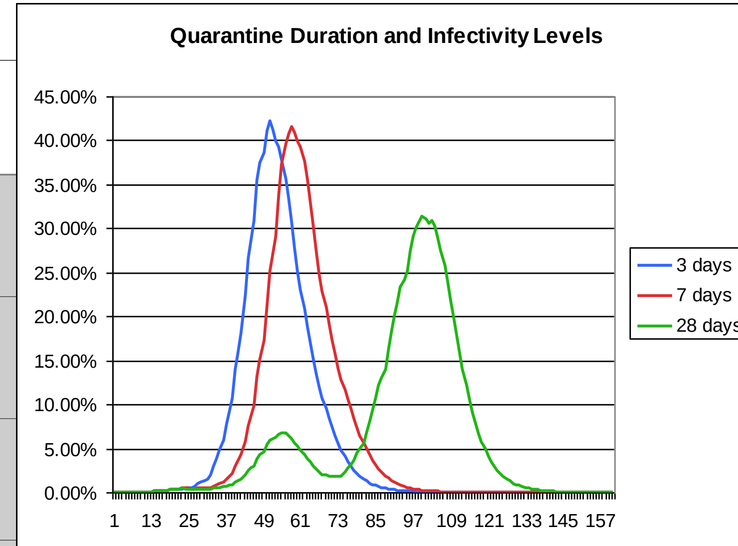
- Automated tools for:
  - Evaluating response policies, data efficacy, attack severity, and detection tools related to weaponized biological attacks *in the presence of background diseases such as flu*
  - Generating high fidelity virtual data for testing detection and fusion algorithms, and exploring potential impact of never before seen events
- Systematically reason about:
  - The rate and spread of disease at the symptom level with high degree of realism
  - Early presentation of diseases as seen in secondary data streams
  - Potential response scenarios, such as inoculation
  - Policy design for potential problems
- Push the frontier of high dimensionality social simulation models (fidelity, precision, speed, comprehensiveness, etc.)

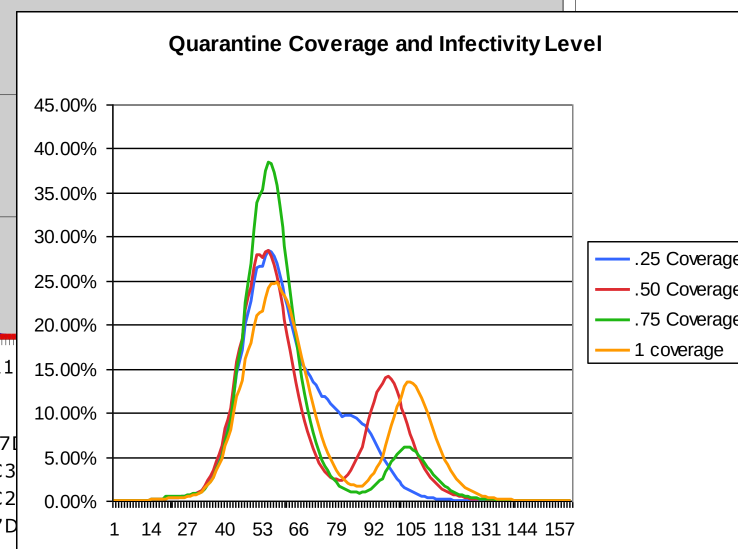
Interventions

- Quarantine
  - Fraction of infected quarantined - .25, .5, .75, 1
  - "reduced contact" - 90% reduction
  - Length – 3 7 28 days
  - When – .05% infected, 1% infected, 2% infected
- School Closures
  - Historic evidence that even when schools close, children are in contact
  - "another Saturday"
  - Length – 3 7 28 days
  - When – .05% infected, 1% infected, 2% infected
- Vaccinations
  - May not be 100% effective - .25, .5, .75
  - May not be 100% distributed - .25, .5, .75

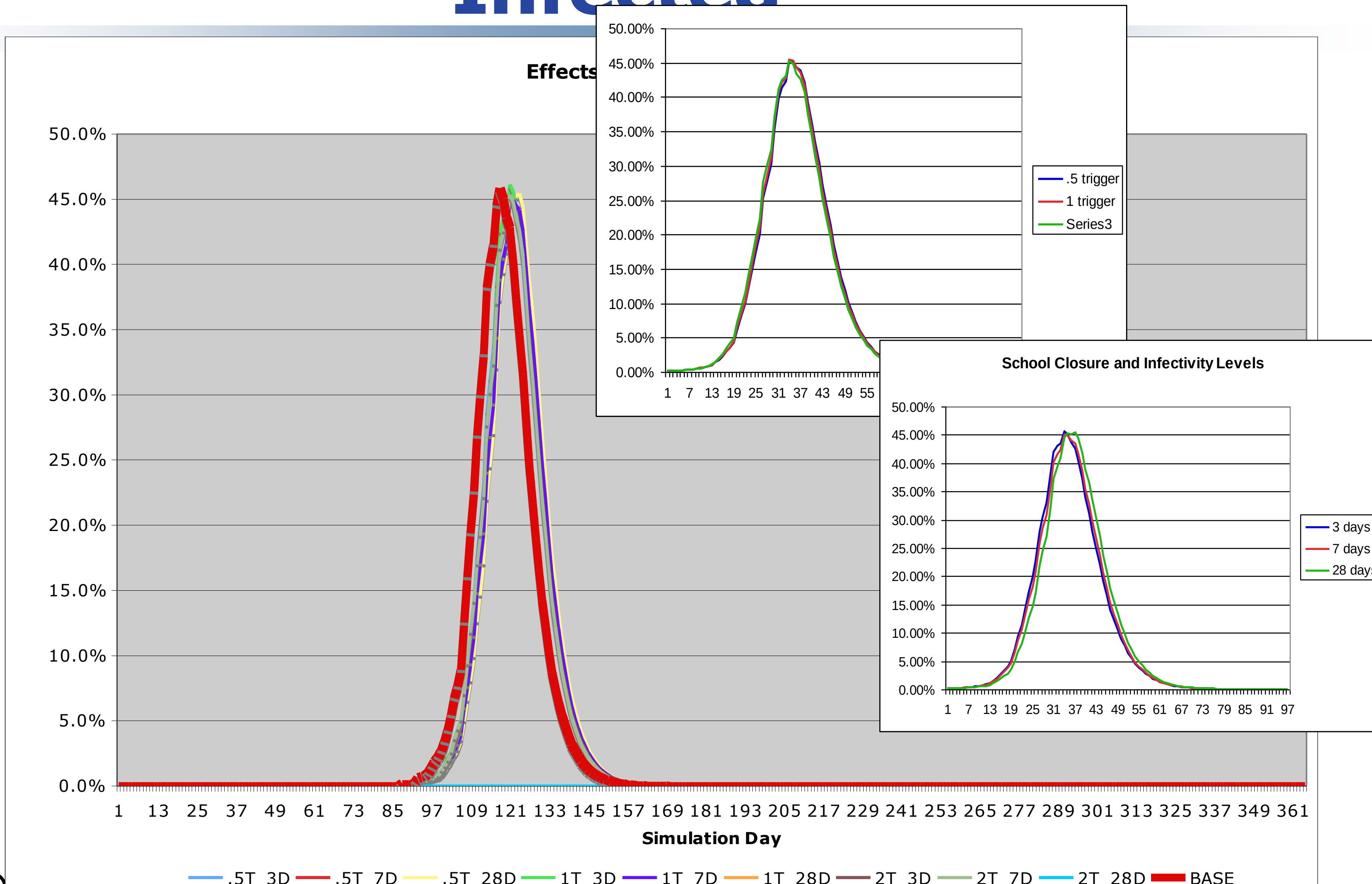
Impact of Quarantine



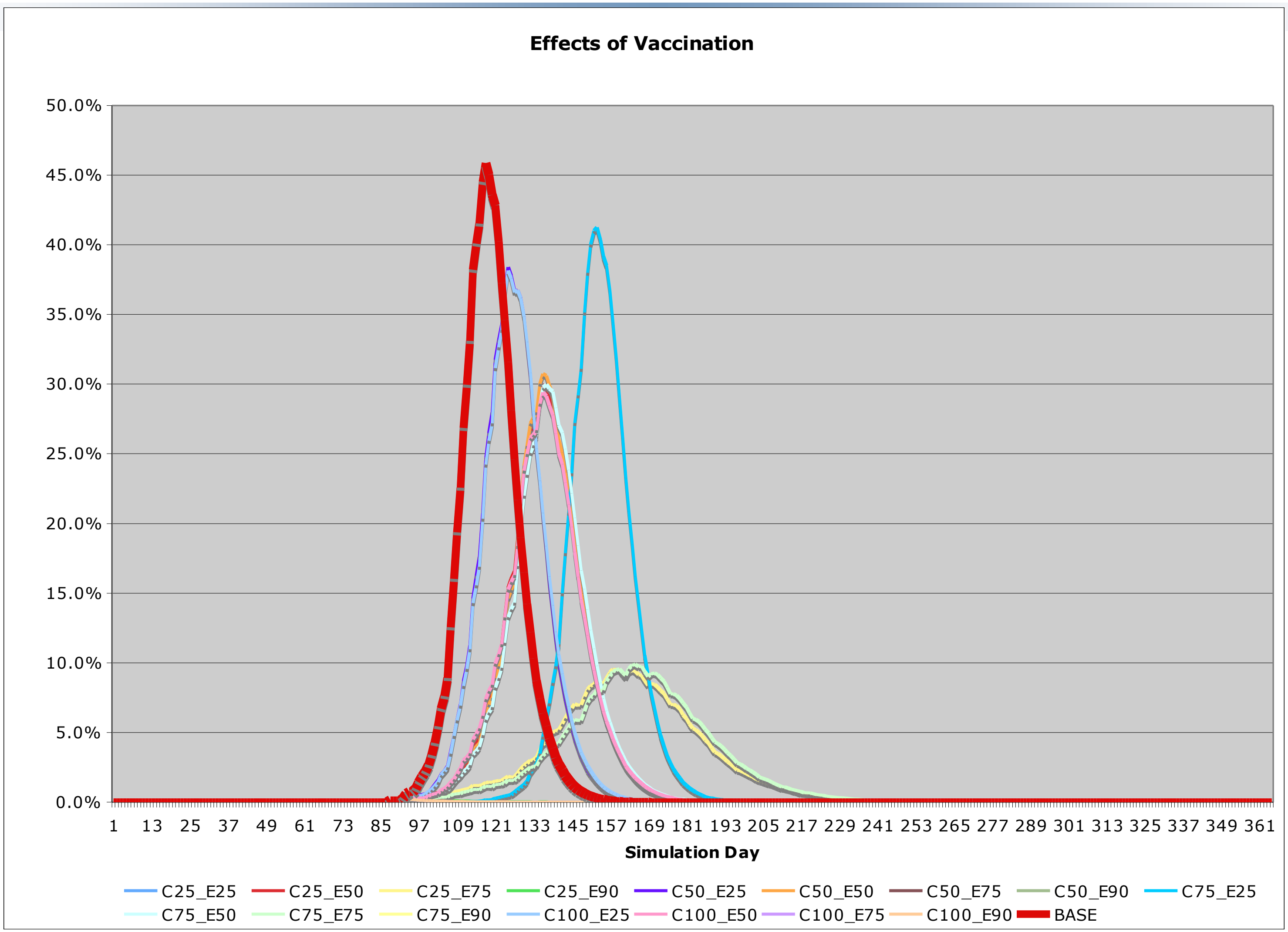




Impact of School Closure on Infected



Impact of Vaccination on Infected



Validated Features

- Anthrax attack & disease model ("docking" or computational model alignment with IPF, Incubation-Prodromal-Fulminant – a revised SIR -- Model)
- Smallpox attack & disease model, docking with SIR
- School absence
- Work absence
- Doctor visit
- ER visit

Validation over Time (C5=last "Challenge", C1..4=previous "Challenges")

Type	C1	C2	C3	C4	C5
Docking: Comparison against another model				✓	✓
Generic Pattern: Showing pattern for each generated data stream matches observed patterns	✓	✓	✓	✓	✓
Characteristic Matching: Showing for each generated output data stream that it has correct seasonal or daily pattern		✓	✓	✓	✓
Relative Timing of Peaks: Showing time between peaks for dif. data streams matches observed dif.		✓	✓	✓	✓
Empirical Pattern: Showing pattern for each generated data stream matches empirical pattern – best for input streams			✓	✓	✓
Within Bounds: Showing for each generated output data stream that the mean of simulated stream falls within min/max of that stream for real data			✓	✓	✓
First moments: Showing for each generated output data stream that mean is not statistically different than real data – yearly, monthly or daily				✓	✓

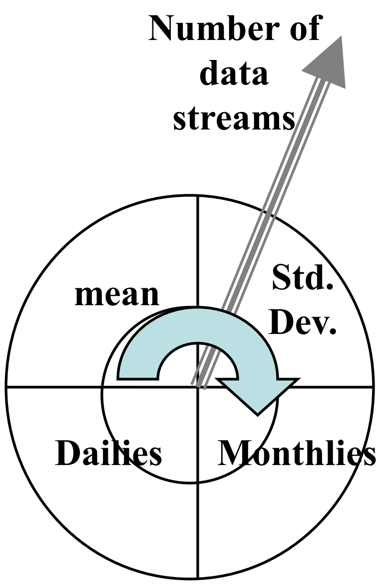
Carnegie Mellon
isr institute for SOFTWARE RESEARCH

Verification & Validation

- Internal Tuning
  - Existing data sets to parameterize
    - Reporting delays
    - Disease profiles
    - Agent social networks
    - Age, race, gender, economic differences on behavior and susceptibility
    - Variation in behavior by time of day, day of week, month, season
    - Usage of IT
  - Sources
    - Behavioral surveys
    - Nursing studies
    - CDC reports
    - Communication studies
    - OTC purchases
  - City profiling
    - Census data
    - School district
    - Maps
- Validation – emergent behavior compared to real data
    - Death reports
  - General behavior
    - Disease replication for historic cases
    - Pharmacy purchases
    - Cold shelf and influenza spike
  - Influenza
    - Grade School Absenteeism
    - ER reports
    - OTC purchases
  - Level
    - General pattern
    - Mean, std
    - Variation in disease reports by day of week, month, season, local

What Data Streams is Validation Done On

Data Stream	C2	C3
Work absenteeism	**Yes**	**Yes**
School absenteeism	**No**	**Yes**
ER visits	**Yes**	**Yes**
Doctor visits	**Yes**	**Yes**
OTC drug purchase	**No**	**Yes**
Sentinel trace	**No**	**No**
Network distribution	**No**	**Yes**



What validation or tuning has been done

- Work absenteeism within the lower & higher empirical bounds
- School absenteeism within the lower & higher empirical bounds
- Doctor visits within the lower & higher empirical bounds
- ER visits within the lower & higher empirical bounds
- Drug sales per group is near the empirical mean
- Face validation of a sentinel population trace
- Automated output check

Sources of Data for Validation

- NCES Indicator 17 & Indicator 42-1, for calculating school absenteeism
- CDC Advance Data, from Vital and Health Statistics, no. 326, 2002, for calculating ER visits
- CDC Advance Data, from Vital and Health Statistics, no. 328, 2002, for calculating doctor visits
- 1997 US Employee Absences by Industry Ranked (http://publicpurpose.com/lm-97absr.htm) for determining work absenteeism
- OTC Sales by Category from AC Nielsen (http://www.chpa-info.org/statistics/otc_sales_by_category.asp) and PSC's FRED data for pharmacy OTC drug sales

CONCLUSION

Why Use Agent-Based Models?

- *Ethical:* Cannot test policies on real populations
- *Preparatory:* Can create hypothetical situations with more potency than existing ones – Can examine wide range of scenarios – Enabling systematic imaginative thinking
- *Cost effective:* Creating new technologies, procedures and legislation for data collection is expensive
- *Faster:* Real time evaluation of existing systems is too time consuming
- *Appropriate:* The world and the simulation are complex non-linear dynamic systems
- *Flexible:* Response to novel situations requires rapid evaluation of previously unexamined alternatives