Carnegie Mellon

Computational Modeling & Analysis

Kathleen M. Carley
Carnegie Mellon University

Center for Computational Analysis of Social and Organizational Systems

http://www.casos.cs.cmu.edu/

Carnegie Mellon

Why Computational Analysis Is Important

- Socio-technical systems are complex
  - Many personnel, resources, tasks, procedures
- Socio-technical systems are dynamic
  - As culture changes structure can change
  - As resources are consumed structure adjusts
- Effects are often non-linear
  - Impact of population growth on culture
  - Impact of stress on individual performance
- Experiments are costly – especially for large groups
- Field studies may be too invasive or dangerous

Under these conditions simulation is valuable

Carnegie Mellon

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

CASOS

Carnegie Mellon

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

CASOS

Carnegie Mellon

50 Years of Computational Theory

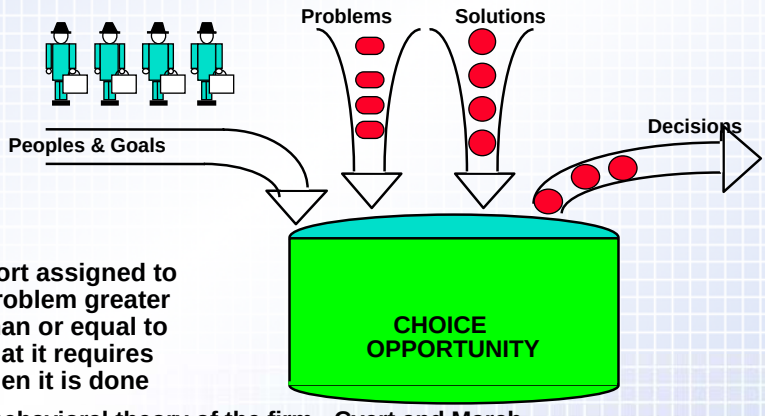
- Computational analysis is revolutionizing social theory
- Important advances in the areas of:
  - Cognition.
  - Interaction.
  - Chance.
  - Adaptation.
- Fundamental results:
  - Bounded rationality.
  - Satisficing.
  - Competency traps.
  - Emergent order.
  - Learning clashes.
- New paradigms such as
  - Neo information processing theory
  - Dynamic network analysis

Question existing paradigms
Explore non-standard assumptions

CASOS

Carnegie Mellon

An Early Model - Garbage Can



If effort assigned to problem greater than or equal to that it requires then it is done

- **A behavioral theory of the firm - Cyert and March**
- **Garbage can model - Cohen, March and Olsen**

CASOS

Carnegie Mellon

A Modern Model - BioWar – conceptualization

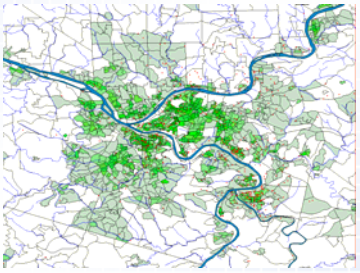
city scale multi-agent network model of weaponized attacks

City Description Manager

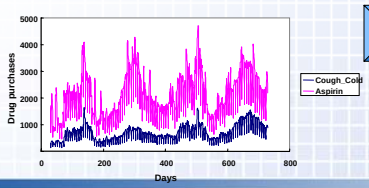
Disease Profile Manager

Attack Scenario Manager

Output Manager



Drug Purchases



Display

Geographic
Chart
All locations
All people
All diseases
Select location
Select subgroup
Select disease
influenza
smallpox

Comparison

Carnegie Mellon

What has changed?

- Bigger, better, faster computers
- Multi-agent modeling and object oriented code
- Specialized simulation languages
- Increased attention to validation
- Incorporation of real data
- Increased standardization of representation schemes
- Increased acceptance in science, engineering and computer science
- Increased divide between types of modeling in part due to purpose of model

Carnegie Mellon

Types of Models

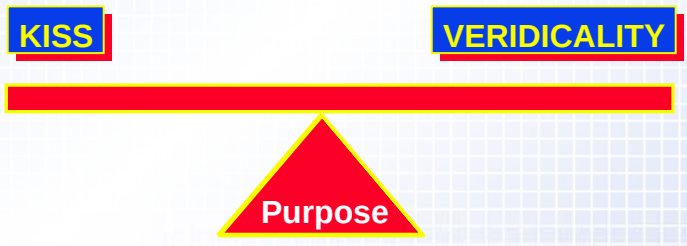
- Intellective -v- emulation
- Stochastic - random number generator (Monte Carlo) -v- deterministic
- Rule based -v- equation
- Learning -v- static -v- optimization
- Centralized -v- multi-agent
- System Dynamic -v- multi-agent -v- multi-agent-network

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

CASOS

Carnegie Mellon

Balance



CASOS

Carnegie Mellon

Kitchen Sink Approach

People buy into model and suggest changes

Ease of adding features

Lack of a minimal theory

Not experiment directed

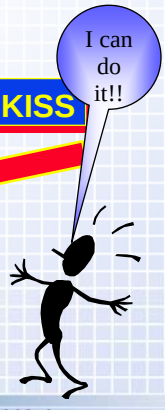
VERIDICALITY

Everything in the model and model everything

characterized by:

large, many parameters, many untested modules

KISS



Purpose

CASOS

June 13, 2006 Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU 11

Carnegie Mellon

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

VERIDICALITY

KISS

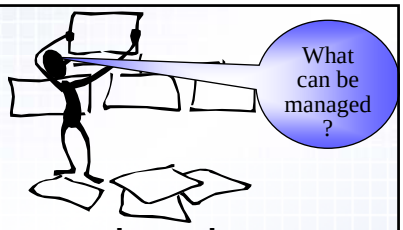
Purpose

CASOS

June 13, 2006 Copyright © 2006 Kathleen M. Carley, CASOS, ISRI, SCS, CMU 12

Carnegie Mellon

Emulative Approach



Built via blocks
Empirically grounded
Addresses applied issues

Sufficient unto the task

characterized by:
large, partly tested modules

KISS

VERIDICALITY

Purpose

Carnegie Mellon

Shifting the Balance

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

KISS VERIDICALITY

Intellective or theory building models

KISS VERIDICALITY

Engineering, emulation, or wind tunnel application models

Carnegie Mellon

General Approach

- Either – Multi-agent models
  - Expect outcomes dependent on actions of large number of heterogeneous actors
  - Realism of agents depends on level of model
- Or – System dynamic models
  - Expect outcomes based on system wide interactions
  - Realism of system depends on purpose of model
- Process intense models
  - Detailed emulation of processes
  - Multiple factors interacting in non-linear ways
- Multi-level models
  - Different learning strategies needed at different response levels (person, organization / institution, nation)
  - Require different kinds of data, analysis, tools at each level
- Empirically grounded
  - Multi-prong approach to real data
  - Moving from face validation to tuning to validation

CASOS

Carnegie Mellon

Multi-Agent – vs – System Dynamic Models

- Multi-agent
  - Building from the ground up
  - Large number of heterogeneous agents
  - Systems put constraints on agents
  - Heuristic based
- System Dynamic
  - Building from the top down
  - No agents
  - Systems are composed of multiple interacting parts
  - Sets of differential or difference equations

CASOS

Carnegie Mellon

Features of Process Intense Models

- Event driven / organized.
- Applicable to multiple problems - knowledge intensive.
- Model is the theory.
- Model does the task.
- Generate large numbers of hypotheses.
- Emulative - Impossible to completely validate.
- Intellective – Unreasonable to validate other than face validation.

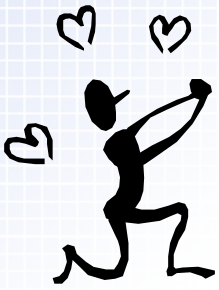
Carnegie Mellon

What should any model of this type do?

- Exhibit emergent behavior
  - At the system or group level there should be behavior that is more than/different than the sum of the parts or the average
- Path dependence
  - Specific outcomes depend on starting point and choices along the way
- Inevitability of change
  - There are many bases for change – learning, evolution, turnover …
- Large number of outcomes
  - Many questions can be addressed with the same model

Carnegie Mellon

OK



I'm sold!!!

Uhm, how do I do this?

Carnegie Mellon

High Level Process for Developing a Computational Model

- Theory or idea
- Identify core entities and relations
- Identify core processes
- Develop a representation scheme
- Identify key question you want to answer
  - What output do you need
  - What do you need to vary
  - Layout one critical virtual experiment
- Go back and reduce everything to just what is needed for experiment
- Code, keeping in mind bigger picture
- Make RP assumptions for details missing in theory

Carnegie Mellon

GULP!!! Can you make it easier?

- Start with an existing model and modify it
- Work in teams
- Use modeling language
- Start simple
- Take a building block approach

Carnegie Mellon

Building Block Approach

Build & Test

Build & Test

Build & Test Build & Test

characterized by:

small, modular, sequence of inter-related virtual experiments, framework

- **Features added by potential for major impact**
- **Focus on necessary components of theory**
- **Focus on experiments**

VERIDICALITY

KISS

Purpose

Carnegie Mellon

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

CASOS

June 13, 2006 23

Carnegie Mellon

Illustration - CONSTRUCT: Building Blocks

NetWatch

DyNet2 | BioWar 2

DyNet | Agentized Construct | BioWar

Construct-TM Transactive memory, referential DB

Construct-O Performance, Triads, new IT

Beliefs | Information Seeking Databases, avatars

Construct – relative similarity People, books

CASOS

June 13, 2006 24

Carnegie Mellon

Basic Construct Model

ACTION

$$Interact_{ij}(t) = f(Availability_i(t), ProbInteract_{ij}(t))$$

$$Communicate_{jik}(t) = f(ProbInteract_{ij}(t), Known_{jk})$$

ADAPTATION

$$Known_{i*}(t+1) = Facts_{i*}(t) + Belief_{i*}(t) + Communicate_{jik}(t)$$

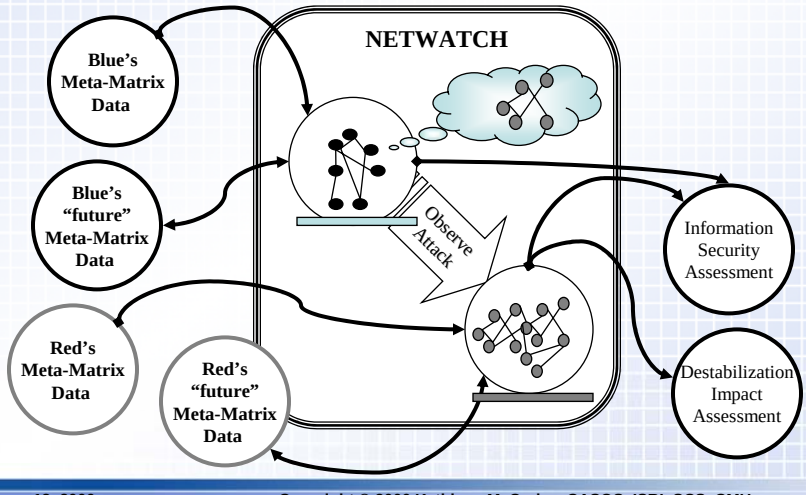
MOTIVATION

$$ProbInteract_{ij}(t) = \frac{SharedFacts_{ij}(t) + SharedBelief_{ij}(t)}{\sum_{h=1}^{I} ShareFacts_{ih}(t) + SharedBelief_{ih}(t)}$$

CASOS

Carnegie Mellon

Advanced: NetWatch:



CASOS

Carnegie Mellon

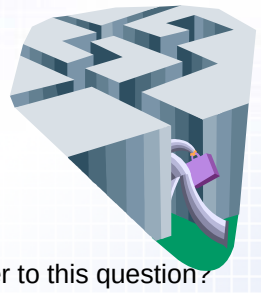
How did we move from here to there?

- Add modules based on
  - What is an interesting question
  - What kind of data can you get
  - What alternative theory do you want to consider
- Utilize virtual experiments
  - Test the system
- Don't assume the results are cumulative
  - These are interacting systems
  - New modules may fundamentally change the results
- Validate
  - What new capability do you have

CASOS

Carnegie Mellon

Getting Started



- What question do you want to address?
- What are the major 2-4 factors?
  - Is there an existing theory?
  - What is the central argument?
- How do these factors interact to impact the answer to this question?
  - Can you draw a picture?
- How will you know if you get this right?
  - Is there an existing model?
  - Is there existing data?
  - Is there a subject matter expert (SME)?
- What kind of data do you need to collect to answer question?
- What kind of parameters do you need to vary to answer question?
- Specify the factor – relate input parameters to output data?

CASOS

