DARPATech, DARPA's 25th Systems and Technology Symposium
August 8, 2007
Anaheim, California
Teleprompter Script for Dr. Sean O'Brien, Program Manager, Information Processing Technology Office

Computational Social Science

» **SEAN O'BRIEN:**
Imagine, that you are the commander leading coalition forces to accomplish an important mission.

You will have to deal with mixed communities and cultures.

You're equipped with a host of resources that includes the traditional military capabilities and funds for projects to help the local populace, but these resources are not unlimited.

As the commander,
you realize that to make effective decisions it is critical that you quickly gain a rich understanding of the complex and dynamic political, cultural, and economic landscape in the area of operations.

In today's wired-in world—you will almost certainly have a glut of information about activities on the ground—from the latest intelligence assessments, right down to live and continuous media feeds on events as they are unfolding.

What you don't have today are really effective decision tools to help you assess the most likely outcomes of the courses of action available to you;

to apply the limited resources in the most effective way to achieve

desired outcomes;

to know the specific factors that are irritants to the communities in the area;

to know how the communities are likely to respond to courses of action undertaken, particularly the infrastructure projects that are started, or stopped;

to know when a course of action taken is actually working in line with expectations;

and to avoid adverse unintended consequences by understanding 2nd and 3rd order effects.

In short, a commander has to answer the fundamental questions: What options do I have at my disposal to meet the mission objectives?

And what are the likely outcomes that would result from executing the different options?

There's a lot of road between here and there.

But taken together,
some recent advances have the potential to move us in the right direction.

First, there is the increasing sophistication and use of **agent-based social simulations.**

Second, we've seen an explosion of new data sources describing communities around the world, who the people are, and what they believe—not just at the nation-state level, but right down to tribes,

neighborhoods, and social groups.

This increase in the quantity and quality of data can dramatically increase the value of agent based simulations.

The third, and probably most important recent development that will make our decision tool possible, is the **proliferation of social, cultural and behavioral theories.**

Many of these theories are in competition.

Most cannot be studied effectively with traditional social science methods.

All of them need to be formalized and integrated.

But the capabilities that IPTO is envisioning may allow us to subject these theories to rigorous empirical evaluations in ways that could revolutionize the manner in which we anticipate and respond to threats from abroad…
and we may revolutionize the social sciences along the way.

I'm going to address each of the three key focus areas in turn.

Agent-based societal simulation environment:

An agent-based model consists of individual agents—representing for example organizations—that are commonly implemented in software as objects.

The agents behave and interact using a simple set of rules.

Running such a model simply involves placing the agents within a

simulation environment, and watching what happens as alternative courses of action are implemented.

Agent-based models have been around for decades and have been applied to study traffic flows and simulate micro
(and macro)
economic behavior, among other uses.

More recently,
we've witnessed an explosion of interest in applying agent-based simulations to study the effects of social and cultural dynamics.

But several things are required to make these simulations operationally useful.

One is the need to create agents that mirror the actual communities and organizations they're designed to represent.

The other is to leverage the hundreds of social, cultural, and behavioral theories to develop a set of rules that govern the interactions among agents.

After all, it is this set of rules that ultimately serves as the basis for anticipating the **reaction** to a commander's actions.

The more closely the rules (or theories) reflect how groups actually behave and respond to events, the more closely the simulated behaviors will mirror real world behaviors.

Understanding the Regional Area

Creating realistic agents requires access to good data on cultures and communities in a regional commander's area of operations… data that

describe a communities' characteristics;
who they are and what they believe about themselves,
their neighbors,
their government,
our own government and policies,
and life in general.

We have long had access to data that describe the features of countries—their type of political structure, aggregate demographic statistics, economic indicators,
and the like.

Significantly, though, we've seen a massive proliferation of real-time global news feeds that provide significant new information.

And we now have tools for turning that unstructured text into structured data that describe the character and intensity of interactions between different groups and organizations.

This has greatly expanded our ability to analyze the dynamic interactions between different communities and the political environment in which they operate.

But several other sources of data have recently expanded the horizon even further.

We can now obtain detailed census reports for nearly every country online-– detailed data on the characteristics of communities throughout the world, distributed by income, education, ethnicity, and religion.

Another example:
The World Values Survey (WVS) was compiled by a global network of social scientists who have surveyed the basic values and beliefs in

more than 80 countries, on all six inhabited continents.

To date,
they have received
92,000 responses
to 250 questions about beliefs... and 25 years worth of data are now available for research.

And there are software systems that automatically measure the intensity of popular support or opposition to nearly any issue in many countries throughout the world,
in multiple languages,
in near-real time,
and on the fly.

Complementing the values survey is the
UN Globe study which collected data on the cultural characteristics in 62 countries, providing insight into how different cultures view the world, how their views shape their standards of behavior, and how they are likely to interact with other cultures.

These data sources can greatly assist in realistically examining communal interactions and responses to events around them.

But having agents that look like real communities tells us ***nothing*** about how they will react under different circumstances.

Today, the outcomes generated by these simulations, though sometimes plausible, have little chance of being reliably predictive in any meaningful sense.

There is a better way….

Which brings me to my third, and perhaps most important development:

Synthesizing and formalizing social, cultural, and behavioral theories

Spread throughout the social science literature are hundreds of theories that provide a set of expectations for how diverse groups, endowed with specific cognitive, demographic, and
cultural characteristics,
develop goals, preferences and standards of behavior;

how they form, alter,
and act upon beliefs;

how they join or leave organizations;

and how they respond to events around them.

Some of these theories have been tested for correspondence
to the real world;

most have not because the tools of the social scientist often do not lend themselves to rigorous empirical evaluations.

And even if they did,
until recently, we lacked the necessary data to perform the analyses.

Here's the challenge:
how do we bring together the most compelling theories, formalize them, integrate them,
and instantiate them within an agent-based environment?

How do we integrate bottom up and top down theories from across multiple disciplines and levels of analysis?

And, how do we make use of the better data sources, and new processing capabilities toward these goals?

And a final challenge:

How do we test the formalized theoretical synthesis for correspondence to
real world behaviors and events?

This is essential to prove or disprove alternative theories, and identify the boundary conditions under which any particular theory or set of theories might apply.

It would also allow us to identify gaps, where new theories would need to be developed to account for discrepancies between simulated and real world behavior.

We need your ideas for new tools and techniques to validate simulated behavior against the real world if we ever hope someday to use this technology to inform confident and reliable decision making.

To be sure, there are some significant technical challenges:
the simulation would need to be quickly populated with real-time data as an operation nears.

Groups and organizations change their repertoires of behavior over time—depending for instance on how much stress or vulnerability they feel.

We would need to come up with new ideas and technologies to identify, process,
and accommodate these changes in a timely manner,

and to make sure the decision tool is dynamic and keeping pace with the changes occurring in the area.

Despite these challenges, we've begun to see a number of automated tools that attempt to analyze unstructured data to provide operational information.

Some of these show promise,
but I challenge you to bring forward your good ideas for even better automated tools.

What if we could do all of this—create a simulation environment, where the agents are designed to represent real cultures and populations of interest—true with respect to their characteristics, beliefs, and standards of behavior—and these agents are observing, processing and reacting to real events as they are reported in current data feeds, and interacting among themselves in empirically realistic ways?

And what if we could confirm that the behaviors we witness in the simulation environment more or less mirror those that are occurring in the real world?

Imagine the possibilities!:

For one, the simulation would provide us with insights on what is occurring in the area of operation...
insights that may not be visible with traditional intelligence tools like HUMINT, COMINT, and SIGINT.

A combatant commander could gain new insight, and develop novel courses of action.

He or she could explore how different strategic, operational, and even

tactical responses could mitigate impending, adverse emergent behaviors and highlight vulnerabilities and unintended consequences — the 2nd and 3rd order effects.

Such a synthetic laboratory could help the commander to make decisions that are much more robust and avoid consequences that are counter-productive to the mission's objectives.

So bring us your ideas, and help us get there.

Please welcome back to the stage Charlie Holland.