

Analysis of C2 and "C2-Lite" Micro-message Communications

AAAI Workshop on Analyzing Microtext

August 8, 2011

Andrew Duchon, Robert McCormack, Brian Riordan, Charlotte Shabarekh, Shawn Weil, Ian Yohai

Contact: aduchon@aptima.com

The views expressed are those of the authors and do not necessarily reflect the official policy or position of the Department of the Army or any other governmental organization.

© 2011, Aptima, Inc.

#### Overview



- Aptima blurb
- Some definitions
- C2 micro-messages
- "C2-Lite" micro-messages
- Common Issues and Data Variation
- General Approach
- Comparisons
- Challenges
- Overall goal: Create a framework such that applications developed for specific datasets can be more easily and broadly applied

## Aptima

- Interdisciplinary small business
  - Founded in 1995
  - 110+ staff (80% graduate degrees)
- Government and commercial clients
  - 350+ contracts with DoD, NASA, DOT...
- Offices
  - Woburn, MA; Washington, DC; Dayton, OH
- Personnel
  - I/O Psychology, Human Factors, Cognitive/Computer/Political Science, ...
- Team Training
  - Understand requirements
  - Training materials to address those requirements
  - Behavioral metrics for measuring performance against those requirements
  - Tools to measure performance
- Human, Social, Cultural, Behavioral Modeling
  - Agent-based modeling of event-triggered opinion change between survey data points
- How can language analysis be applied to automate this work?



### Definitions...



- C2: Command and Control
  - "The exercise of authority and direction by a properly designated commander over assigned and attached forces in the accomplishment of the mission"
  - Hierarchical organizations (teams of teams ...)
    - Information flows up
    - Orders flow down
    - Coordination between and within
- C2-Lite: {Circulation, Comprehension, Coordination, Cooperation, enCouragement} choose 2
  - Same overall goals of C2, but with less authority and less recourse
  - Self-organizing (groups and groups and individuals...)
    - Suggestions are made
    - Information spreads
    - Self-interested cooperation

#### Micro-messages

- Any type of media
  - Must be capable of being persisted
  - Just about everything these days can be...
- Any semi-directed communication
  - At least has a sender, might be broadcast
- Might have content
  - If it doesn't, it needs a receiver
  - Something to link the entities involved
- Really should have a timestamp...
  - Otherwise too many assumptions could be made

#### C2 Communications Analysis: To assess training and operational performance



Performance -

Collaboration

Communication



#### How do we reverse the arrows?

#### C2-Lite Communication Analysis: To understand and ask better questions



Understanding

Prediction



#### Interaction



#### How do we reverse the arrows?

#### **Common Issues**



- What are they talking about?
  - General topic: sports, politics, offense, defense
  - Specific topic: protests, bombings, water distribution
  - Information extraction: person, place, thing, time, mission
- How are they talking about it?
  - Language/dialect
  - Sentiment/valence
  - Dialogue acts
  - Discourse/process structure
- How are they interacting?
  - Directed communications network
  - What/who they are actually responding to
  - Interaction dynamics/stability
- How to understand and assess with fewer experts?
  - Specific enough to be meaningful
  - General enough to be used in time and see the big picture
- What is the best way to relate it to other data?
  - Measures of performance
  - Measures of behavior

#### **Common Data Covariation**



- Size of the data set
  - From: dozens messages from a few people
  - To: hundreds of messages from scores of people
  - To: thousands of messages from millions of people
- Breadth of discussion
  - From: a single task being discussed
  - To: multiple, interleaved tasks
  - To: almost anything
- Protocol
  - From: strict protocol captured by a finite-state machine
  - To: "typical" written or spoken text
  - To: asyntactic stream of consciousness <140 characters</li>

#### **General Approach**





#### C2 Message Analysis



- Radio Communications Between F16 Pilots and an AWACS controller
  - Goal: automate individual and team performance assessment
  - Content: ruled-based normalization given limited vocabulary and protocol
  - Network: information added as additional features (sender sequence)
  - Chat in a Dynamic Targeting Cell
    - Goal: automate individual and team performance assessment
    - Content: addressee extraction, valence, dialogue acts, threading
    - Network: centrality of commander in sender-sequence network
    - Multimodal Communications in an Army Exercise
      - Goal: automate measurement of team-theoretical constructs
      - Content: topic modeling, dialogue acts, valence, threading, etc.
      - Network: distance in communication space to the commander

3.0\*

 $\bigtriangledown$ 

Ö

### **C2-Lite Message Analysis**





#### SALUTE Reports

- Goal: Understand who is who from tidbits of information
- Content: lightly annotated entities (people), verbs (talks, seen with), attributes (clothing, hair)
- Network: Use Infinite Relational Modeling (IRM) to simultaneously cluster entity references and attributes, or entities and relations
- APAN Haiti HA/DR
  - Goal: Understand who is talking about what and in which forum during relief efforts
  - Content: Topic modeling of message contents characterizes rooms and people
  - Network: See the evolution from core users in all forums on more general topics, to usage by just a few users on specific topics
- Social Media
  - Goal: Understand what is happening and where it might lead
  - Content: topic modeling, dialogue acts, info extraction, valence, threading
  - Network: Epidemiological modeling of "memes" to determine which ideas are spreading/will spread

# Shared Characteristics and Analyses



Characteristics	Analyses	C2	C2-Lite
<ul><li>Limited communicators</li><li>Fairly focused</li><li>Ostensible protocol</li></ul>	<ul> <li>Information Extraction</li> <li>Person-Center/Group</li></ul>	Pilot	SALUTE
	network/timing	Radio	Reports
<ul> <li>Dozens of communicators</li> <li>Interleaved, focused</li></ul>	<ul> <li>Specific Topics</li> <li>Dialogue Acts</li> <li>Person-Room</li></ul>	DTC Chat	APAN
topics <li>Natural syntax</li>	network threading		Forum
<ul> <li>100s-1000s Communicators</li> <li>Unfocused, but event- cued</li> <li>Compressed or unavailable content</li> </ul>	<ul> <li>General Topics</li> <li>Spread of sentiment</li> <li>All types of networks</li> </ul>	Multi- modal Comms	Social Media

#### Challenges



- Data from different media
- Time synchronization across media
- The right level of detail
  - For analysis
  - For presentation
- Apply rougher, statistical methods to isolate the messages of real interest for finer, domain-specific knowledge
- Assess C2-Lite situations as if the were true C2
- Find C2 environments behaving like C2-Lite

## Thanks! Questions?



- Acknowledgments
  - AFOSR #FA9550-07-C-0134,
  - L-3 Communications #PO-DS11223
  - ONR (N00014-10-M-0453, N00014-08-C-0188, N00014-09-M-0189)
  - ARI (W91WAW-09-C-0058).
- The views expressed are solely those of the authors.

