U.S. ARMY COMBAT CAPABILITIES DEVELOPMENT COMMAND – ARMY RESEARCH LABORATORY

Algorithmically identifying strategies in multi-agent game-theoretic environments

Erin Zaroukian

Cognitive Scientist

Social Terrain Modeling Team, Multilingual Computing and Analytics Branch, Computational and Information Sciences Directorate

17 APR 2019

APPROVED FOR PUBLIC RELEASE

INTRODUCTION

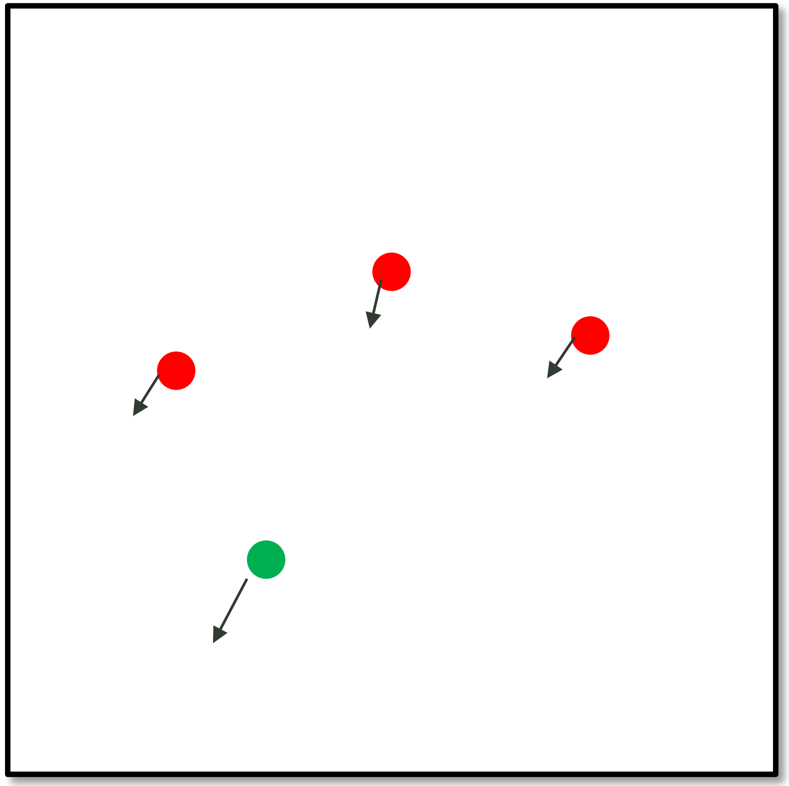
- Computational agents should support their human teammates by adapting their behavior to the humans' strategy for a given task in order to facilitate mutually-adaptive behavior within the team.
- While there are situations where human strategies are top-down, explicit, and easy to understand, human strategies are often implicit and *ad hoc*.

- Our goal: Identify and label the implicit human strategies
  → Facilitate transparency, promote trust, and provide a better understanding of how humans work together and how computational teammates can be trained to fit into a human-human dynamic.

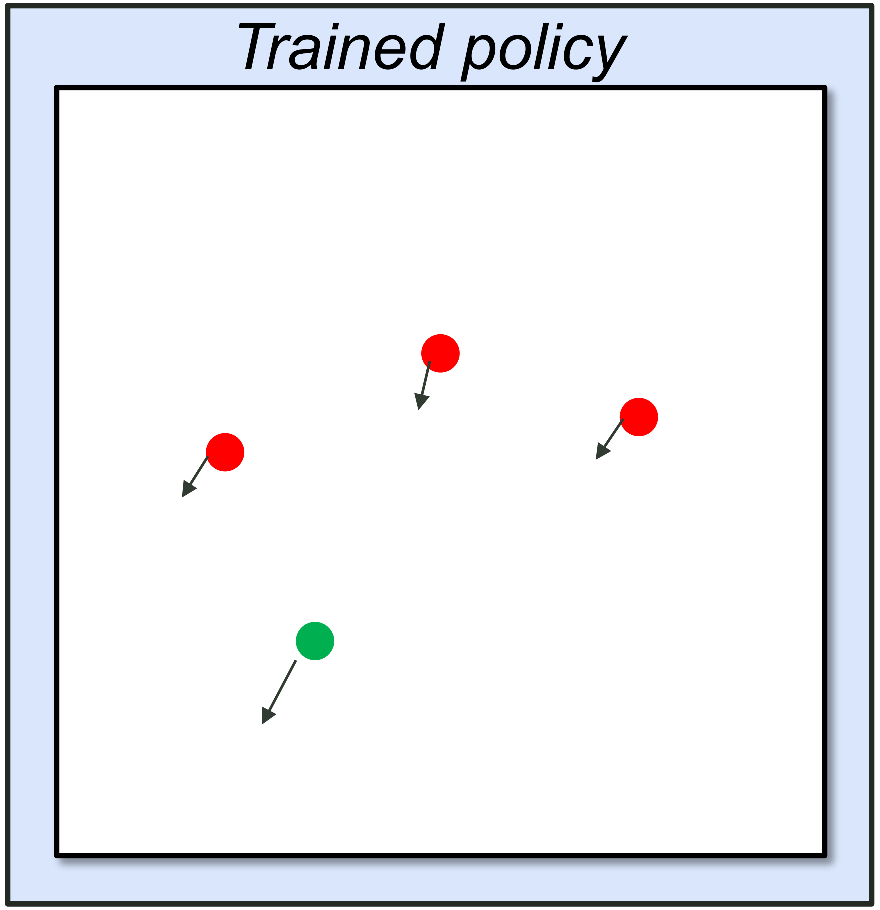
STRATEGIES IN MOVEMENT DATA

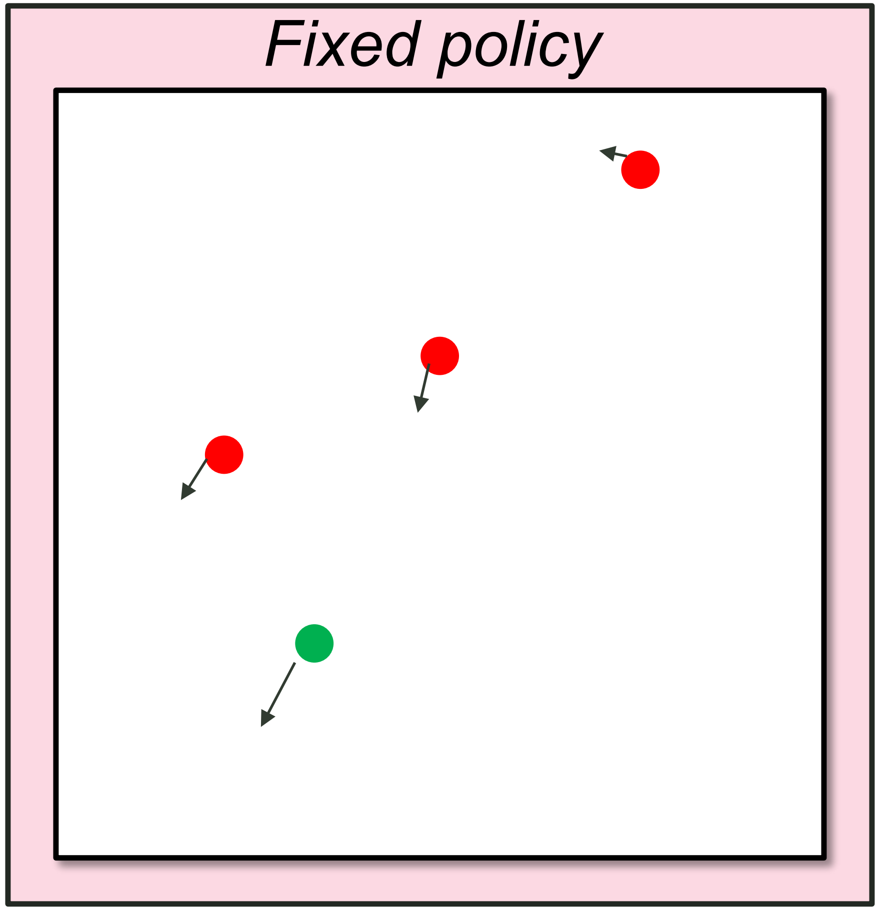
- **Strategies aren't observable! Infer through measurements of behaviors toward a goal.**
- **Existing methods for identifying strategies often require:**
  - Verbal reports of strategy
  - *A priori* set of strategies to recognize
    - e.g., REIative MOtion
  - *A priori* chunking / atomic units of movement data (usually in highly constrained environments)
    - e.g., Context Free Grammars, Linear Temporal Logic
  - Repetition
    - e.g., ALCAMP
- **We use timeseries techniques**
  - Univariate measure of group configuration – *polygon area*
  - Identify strategies through Change Point Detection (CPD) and Dynamic Time Warping (DTW)

METHOD – PREDATOR-PREY PURSUIT ENVIRONMENT



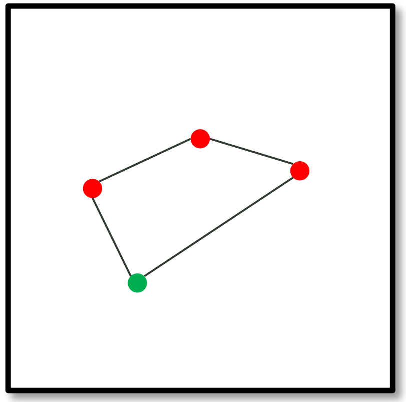
METHOD – PREDATOR-PREY PURSUIT “STRATEGIES”



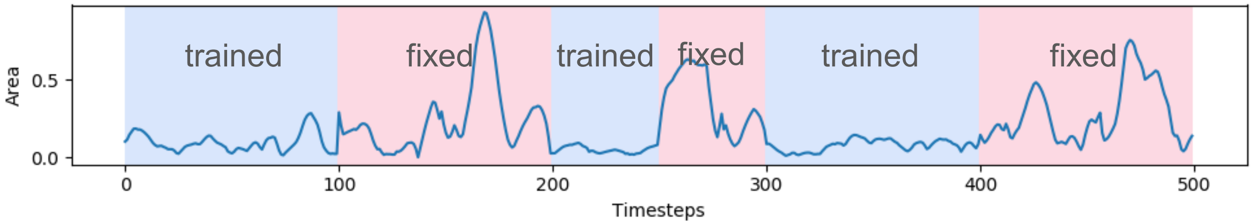


METHOD - TIMESERIES

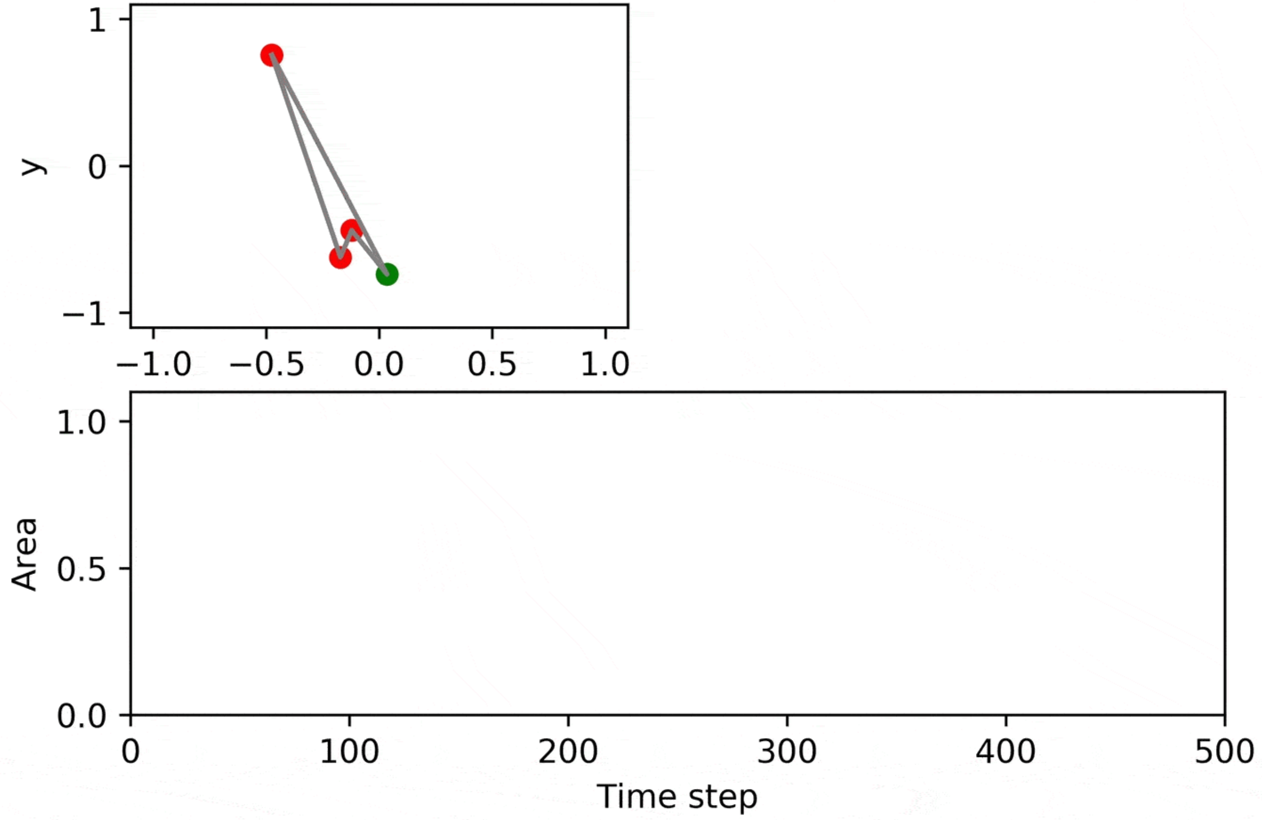
- **Polygon area**



- **Timeseries**
  – Test episodes were creating by interleaving different strategies, i.e., *ground truth segments*

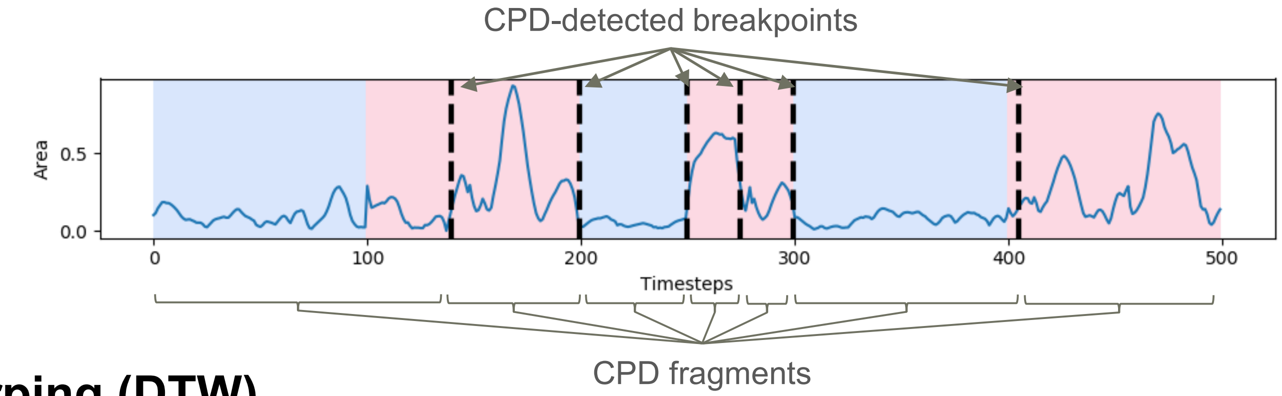


EXAMPLE – EPISODE 2

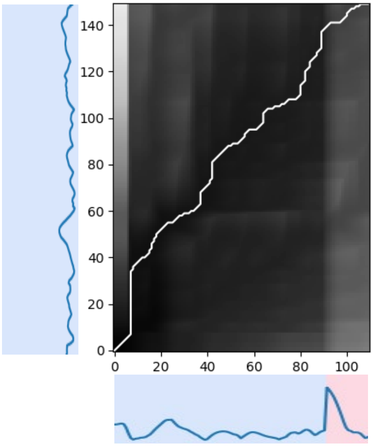


METHOD – STRATEGY IDENTIFICATION

- **Change Point Detection (CPD)**
  - Combination of various cost functions (mean, variance, covariance, rank, density, etc.) from different distributions of data was utilized to determine change points in the timeseries
    - Divide data into *CPD fragments*, which can be compared to ground truth segments



- **Dynamic Time Warping (DTW)**
  - Compare similarities between pairs of timeseries (CPD fragments).

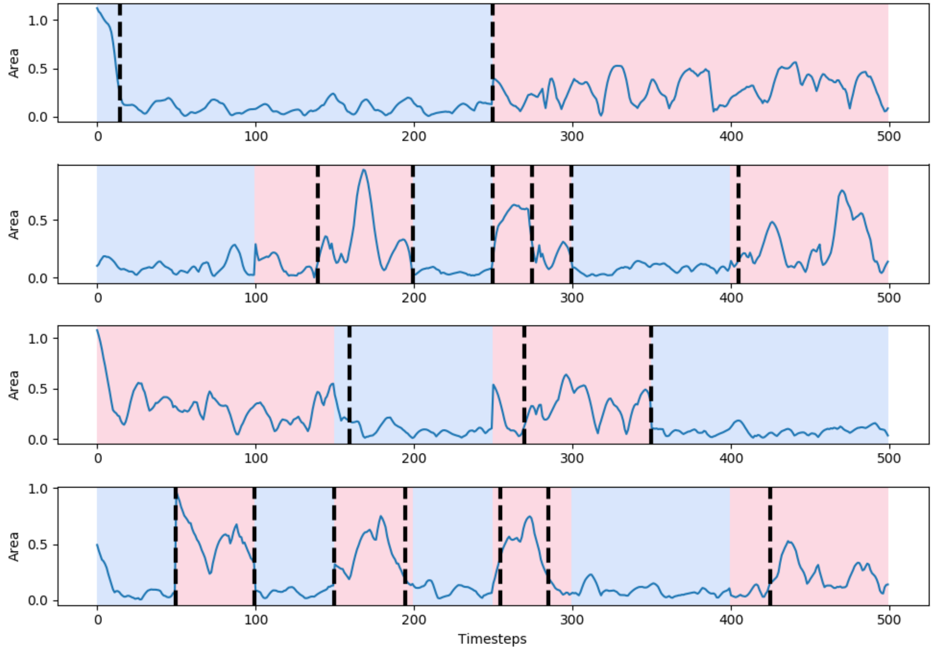


DTW similarity scores

Fragment	2	3	4	5	6	7
1		X				
2						
3						
4						
5						
6						

RESULTS

- **CPD**
  - Metrics comparing ground truth to CPD breakpoints

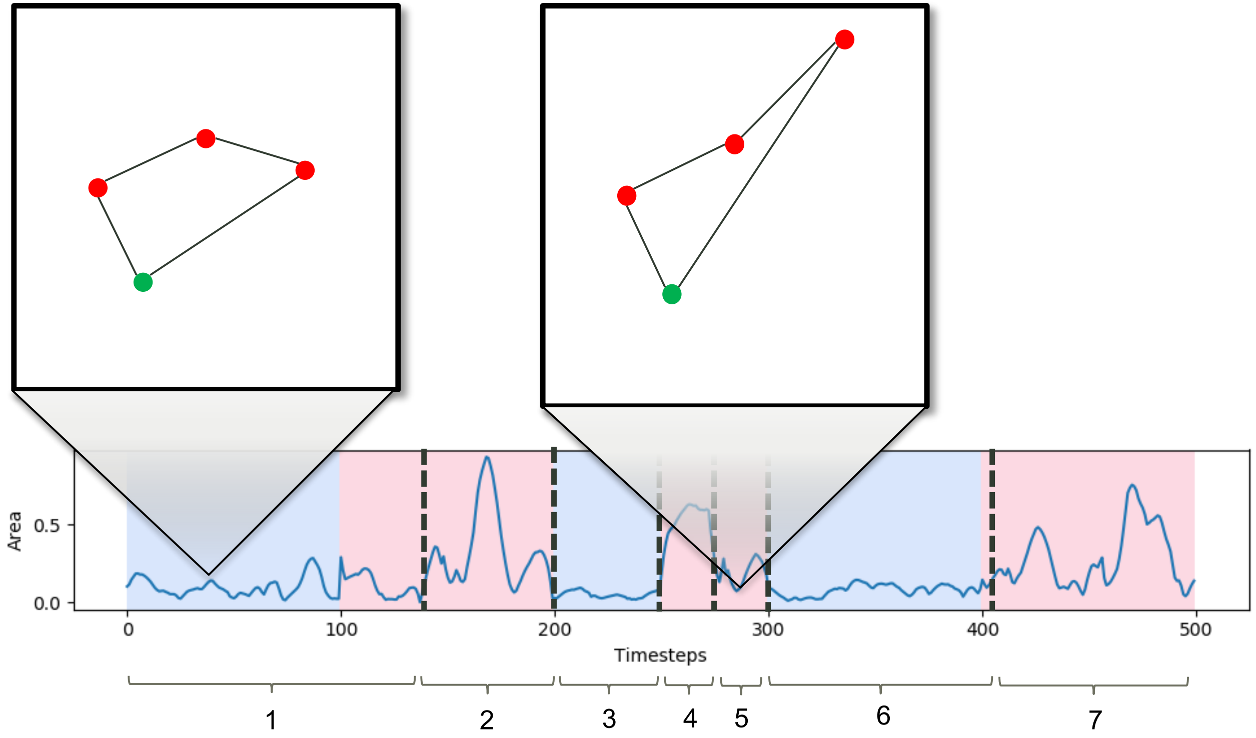


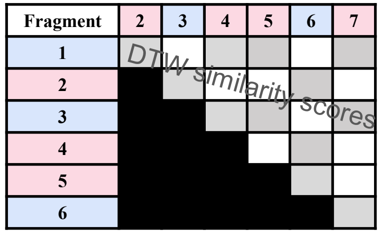
	Precision, Recall 20 timestep margin	Rand index	Hausdorff distance (/500)
Episode 1	0.50, 1.0	0.97	235 (47%)
Episode 2	0.67, 0.80	0.94	40 (8%)
Episode 3	0.67, 0.67	0.95	20 (4%)
Episode 4	0.85, 0.85	0.94	25 (5%)

- **DTW**
  - Similarity scores between same-strategy CPD fragments (median = 0.96) > different-strategy CPD fragments (median = 0.90, Mann-Whitney U = 675, n = 58, $p < 0.001$, r = 0.55).

DISCUSSION

- **Our goal:** Identify and label the implicit human strategies → Facilitate transparency, promote trust, and provide a better understanding of how humans work together and how computational teammates can be trained to fit into a human-human dynamic.
  - Using timeseries techniques, multi-agent predator-prey pursuit task and policy as ground truth
    - Represent group configuration as polygon area
    - Segment with CPD
    - Classify with DTW





DISCUSSION

- **Limitations**
  - How well will this method work
    - When obstacles are introduced to the predator-prey environment?
    - When human teammates are introduced?
  - What information goes into the timeseries
    - Polygon area looses information, try dynamic factor analysis
    - May depend on strategies
  - CPD
    - Cost
    - Sampling rate / quantity of data
  - DTW
    - High similarity for *different* strategies!
  - If strategies are unobservable, how useful is comparison to “ground truth” (policies)?

- **Extensions**
  - t-distributed stochastic neighbor embedding (TSNE)
    - How specific behaviors are linked to the activations of the network → Do strategy/policy changes map to changes in NN activation?
  - Information Theoretic Disentanglement
    - Can strategy be disentangled via deep NN?

SLIDE TITLE GOES HERE

- **Thanks to Sebastian S. Rodriguez, Sean L. Barton, James A. Schaffer, Brandon Perelman, Nicholas R. Waytowich, Blaine Hoffman, Derrik E. Asher, and Jonathan Z. Bakdash**

APPROVED FOR PUBLIC RELEASE

DTW DISTANCE AND SIMILARITY SCORES

Episode 1 Fragment	2	3
1	0.19 **(0.81)**	0.14 **(0.87)**
2		0.07 **(0.93)**

Episode 2 Fragment	2	3	4	5	6	7
1	0.07 **(0.93)**	0.04 **(0.97)**	0.14 **(0.86)**	0.04 **(0.96)**	0.02 **(0.98)**	0.05 **(0.95)**
2		0.14 **(0.86)**	0.08 **(0.92)**	0.09 **(0.91)**	0.10 **(0.90)**	0.03 **(0.97)**
3			0.21 **(0.79)**	0.06 **(0.94)**	0.01 **(0.99)**	0.13 **(0.87)**
4				0.14 **(0.86)**	0.16 **(0.84)**	0.08 **(0.93)**
5					0.04 **(0.97)**	0.08 **(0.92)**
6						0.09 **(0.91)**

Episode 3 Fragment	2	3	4
1	0.09 **(0.91)**	0.05 **(0.95)**	0.10 **(0.90)**
2		0.08 **(0.92)**	0.01 **(0.99)**
3			0.09 **(0.91)**

Episode 4 Fragment	2	3	4	5	6	7	8
1	0.18 **(0.82)**	0.03 **(0.97)**	0.12 **(0.88)**	0.03 **(0.97)**	0.11 **(0.89)**	0.02 **(0.98)**	0.05 **(0.95)**
2		0.24 **(0.76)**	0.08 **(0.92)**	0.18 **(0.82)**	0.07 **(0.93)**	0.195 **(0.81)**	0.12 **(0.88)**
3			0.18 **(0.82)**	0.02 **(0.98)**	0.21 **(0.79)**	0.02 **(0.98)**	0.09 **(0.91)**
4				0.11 **(0.89)**	0.03 **(0.97)**	0.14 **(0.86)**	0.05 **(0.95)**
5					0.18 **(0.82)**	0.02 **(0.98)**	0.06 **(0.94)**
6						0.17 **(0.83)**	0.07 **(0.93)**
7							0.04 **(0.96)**

APPROVED FOR PUBLIC RELEASE