Multi-Agent Learning

by Joseph Johnson

"Intelligence is deeply and inevitably coupled with interaction".

Gerhard Weiβ

Unwelcomed Guest: NFL

What are the challenging problems in Data Mining

- Contributors
  - Pedro Domingos
  - Charles Elkan
  - Johannes Gehrke
  - Jiawei Han
  - David Heckerman
  - Daniel Keim
  - Jiming Liu
  - David Madigan
  - Gregory Piatetsky-Shapiro
  - Vijay V. Raghavan
  - Rajeev Rastogi
  - Salvatore J. Stolfo
  - Alexander Tuzhiliin
  - Benjamin W. Wah

10 Problems in DM

1. Developing a Unifying Theory of Data Mining
2. Scaling Up for High Dimensional Data and High Speed Data Streams
3. Mining Sequence Data and Time Series Data
4. Mining Complex Knowledge from Complex Data
5. Data Mining in a Network Setting
6. Distributed Data Mining and Mining Multi-agent Data
7. Data Mining for Biological and Environmental Problems
8. Data Mining-Process Related Problems
9. Security, Privacy and Data Integrity
10. Dealing with Non-static, Unbalanced and Cost-sensitive Data

Andre Weil

"The great mathematician of the future, as of the past, will flee the well-trodden path. It is by unexpected rapprochements, which our imagination would not have known how to arrive at, that he/she will solve, in giving them another twist, the great problems which we shall bequeath to him/her."

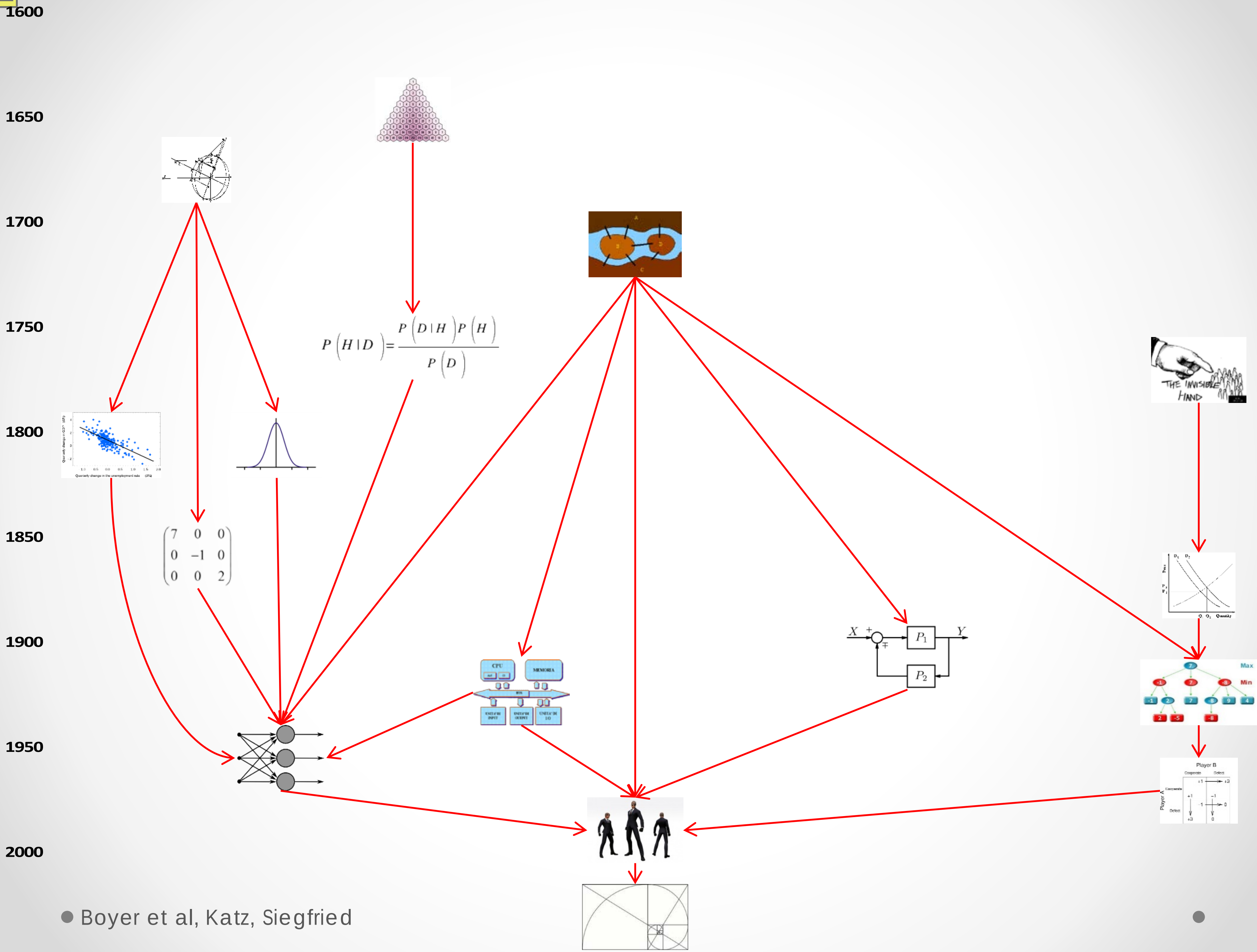
"In the future, as in the past, the great ideas must be simplifying ideas."

- Boyer et al

David Hilbert

"[I]n mathematical science… every real advance goes hand in hand with the invention of sharper tools and simpler methods which at the same time assist in understanding earlier theories and cast aside older more complicated developments."

a talk at the International Congress of Mathematicians in Paris on 8 August 1900, found on BrainyQoutes.com



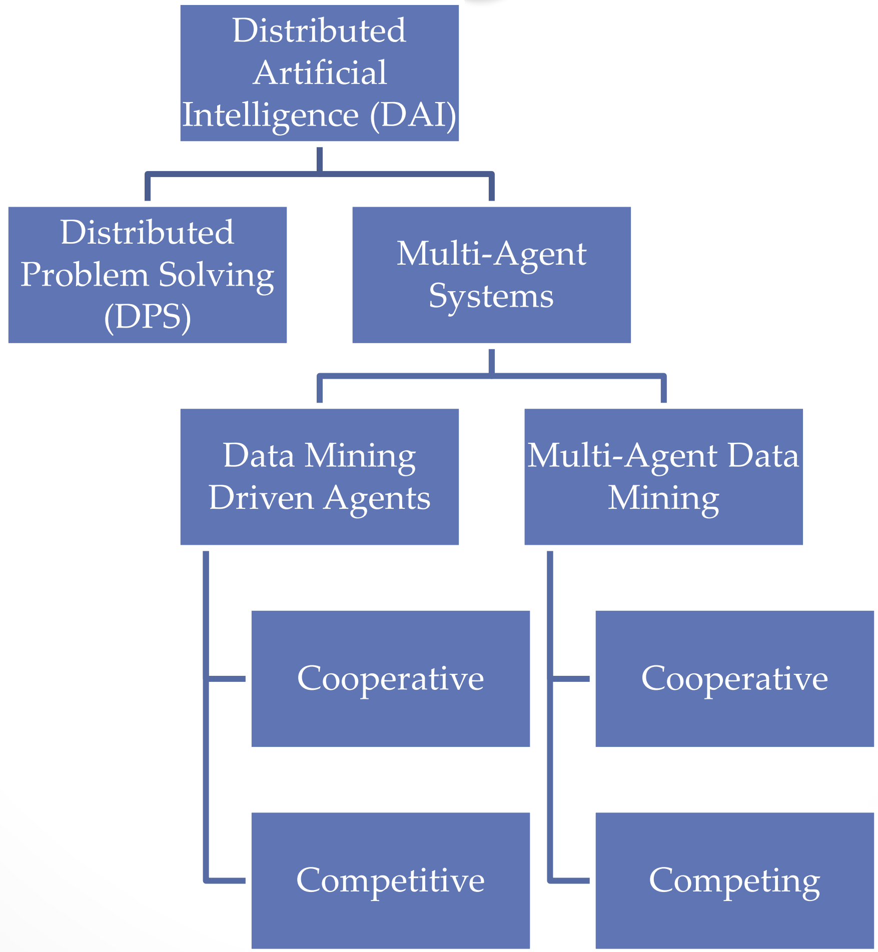
- Boyer et al, Katz, Siegfried

3 Most Common Results in Literature on MAL

1. Convergence of the strategy profile of an (e.g. Nash) equilibrium
    1. Q-learning
2. Successful learning of an opponent’s strategy (or opponents’ strategies)
    1. Rational learning
3. Obtaining payoffs that exceed a specified threshold
    1. No-regret learning

- Shoham et al ,"If multi-agent learning is the answer, what is the question?"

Distributed Artificial Intelligence



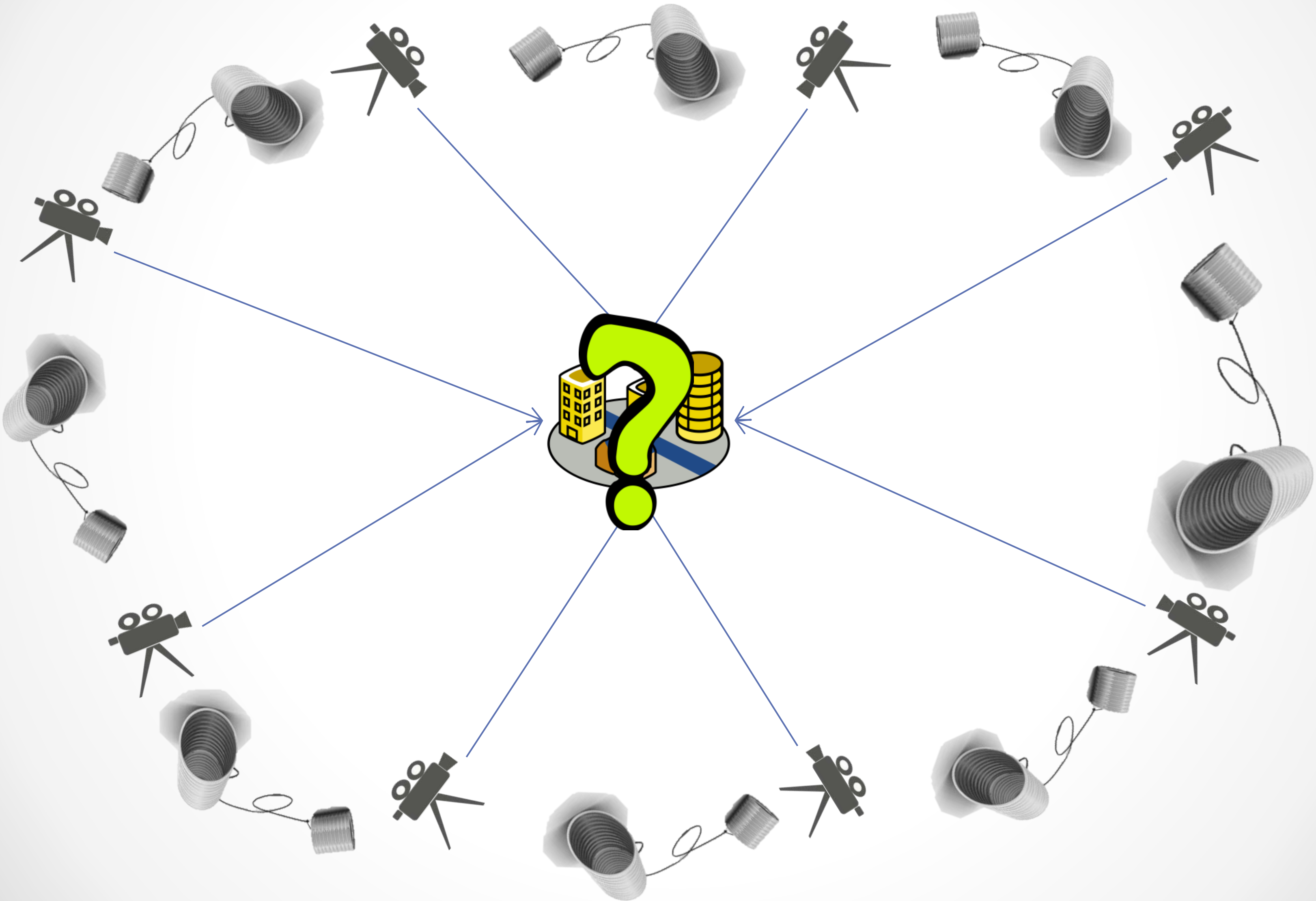
- Chaimantree et al, Panait et al, Stone et al

Multi-Agent Systems

- Agent:
  - "A computational mechanism that exhibits a high degree of autonomy, performing actions in its environment based on information (sensors, feedback) received from the environment."
  - Possesses goals, actions, domain knowledge.
- Multi-Agent Environment:
  - More than one agent
  - Heterogeneity v Homogeneity
  - Constraints on the information agents can sense
  - Communication

10 Problems in DM

1. Developing a Unifying Theory of Data Mining
2. Scaling Up for High Dimensional Data and High Speed Data Streams
3. Mining Sequence Data and Time Series Data
4. Mining Complex Knowledge from Complex Data
5. Data Mining in a Network Setting
6. Distributed Data Mining and Mining Multi-agent Data
7. Data Mining for Biological and Environmental Problems
8. Data Mining-Process Related Problems
9. Security, Privacy and Data Integrity
10. Dealing with Non-static, Unbalanced and Cost-sensitive Data



Scaling Up for High Dimensional Data and High Speed Data Streams

- Data streams in extremely large databases
  - Need for a distributed system to handle the streams
  - "concept drift", "environmental drift"
    - Need to update models while streaming
- Paradox
  - More data allows for use of more complex classifiers
  - However, simpler learners are used in practice because of training time for complex classifiers
  - Need faster ways to learn complex classifiers

An example of DPS: LRTA*

procedure LRTA*

$i \leftarrow s$

while i *is not a goal node* **do**

  foreach *neighbor* j **do**

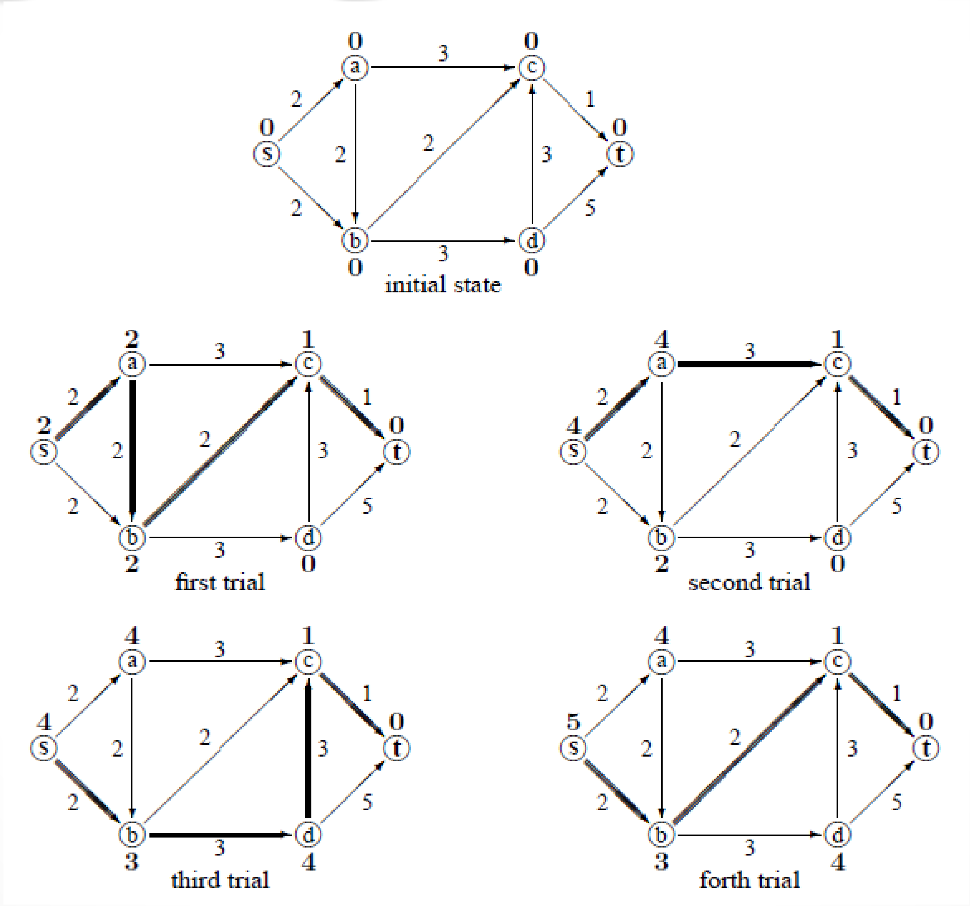
    $f(j) \leftarrow w(i, j) + h(j)$

  $i' \leftarrow \arg\min_j f(j)$

  $h(i) \leftarrow \max(h(i), f(i'))$

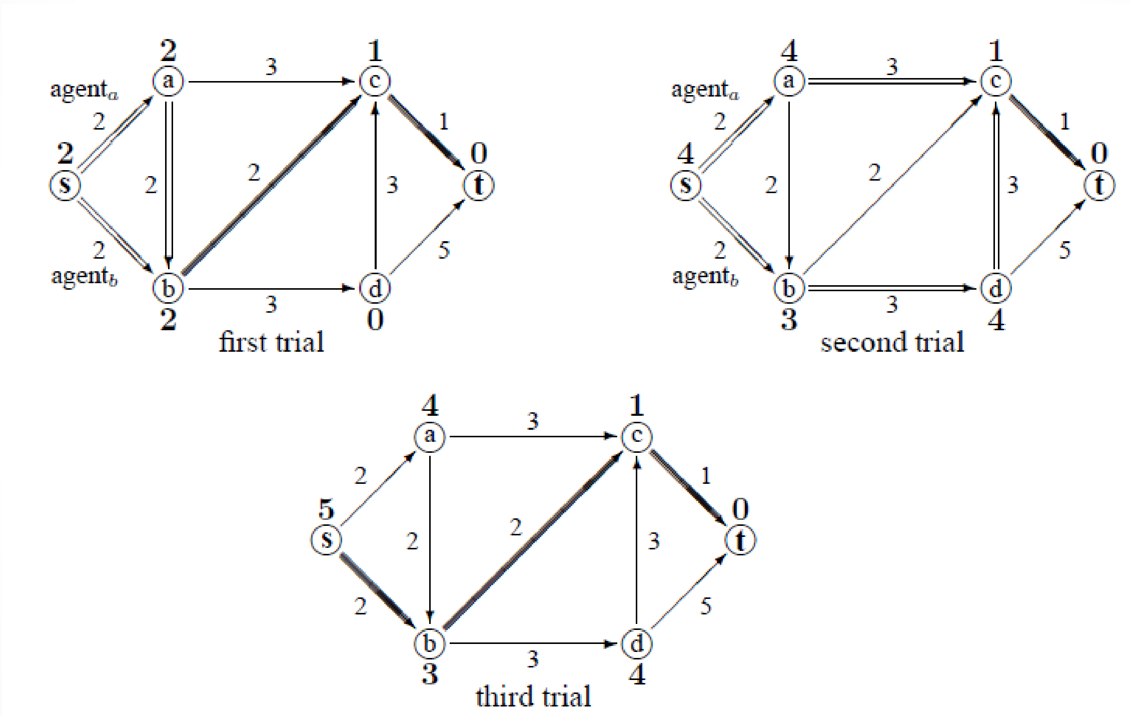
  $i \leftarrow i'$

Single-Agent Example



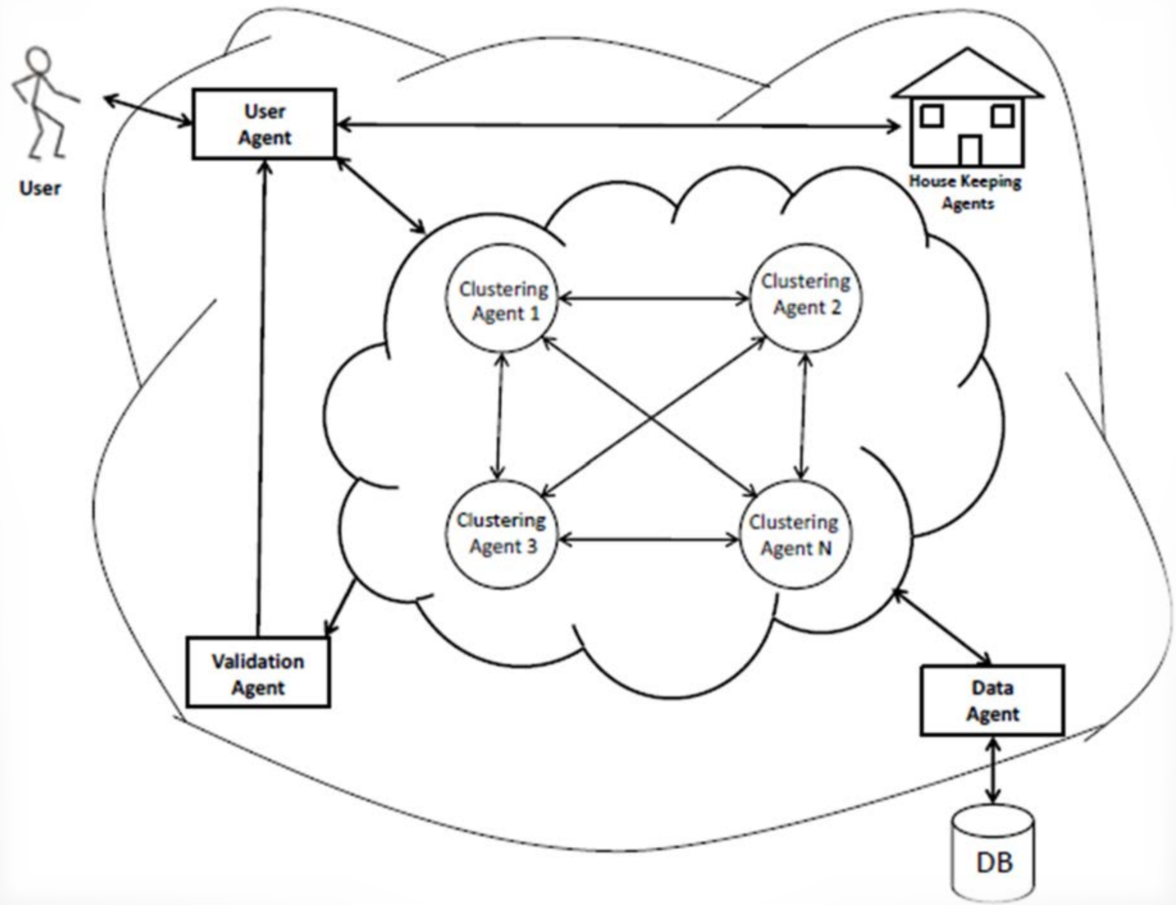
• Shoham et al, *Multi-Agent Systems*

Multi-Agent Example



Shoham et al, *Multi-Agent Systems*

Multi-Agent Clustering

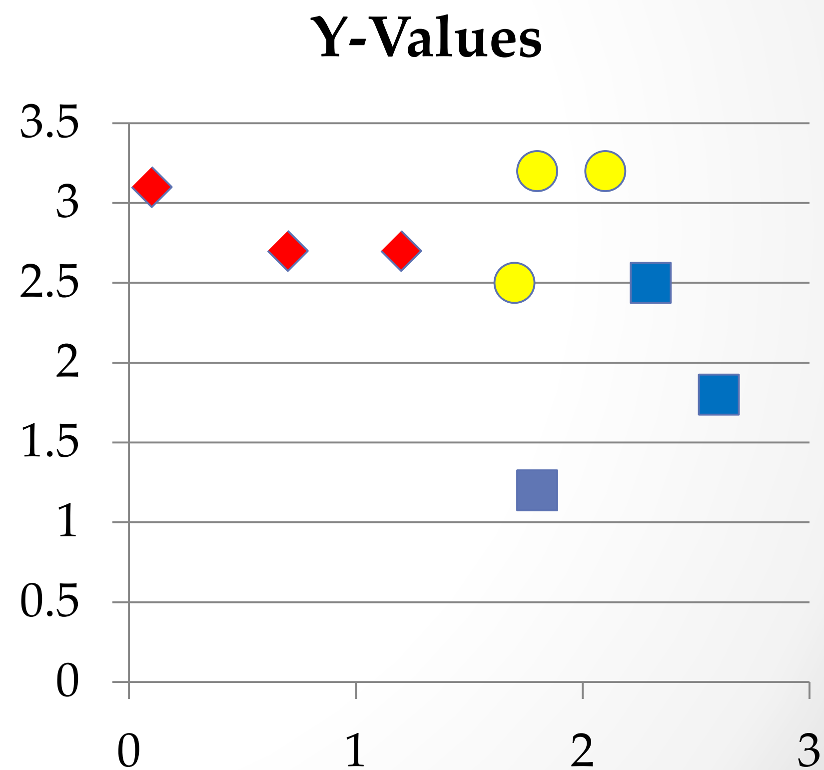


Multi-Agent *k*-Means

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	?
0.7	2.7	Red	?
2.3	2.5	Blue	?
1.8	3.2	Yellow	?
2.6	1.8	Blue	?
0.1	3.1	Red	?
1.2	2.7	Red	?
2.1	3.2	Yellow	?
1.8	1.2	Blue	?

Graphical Depiction



Bidding Phase

Phase I: Bidding using K-means

Input: Dataset ($D = \{d_1, d_2, \cdots, d_n\}$), the desired number of clusters (K)
Output: An initial clustering configuration

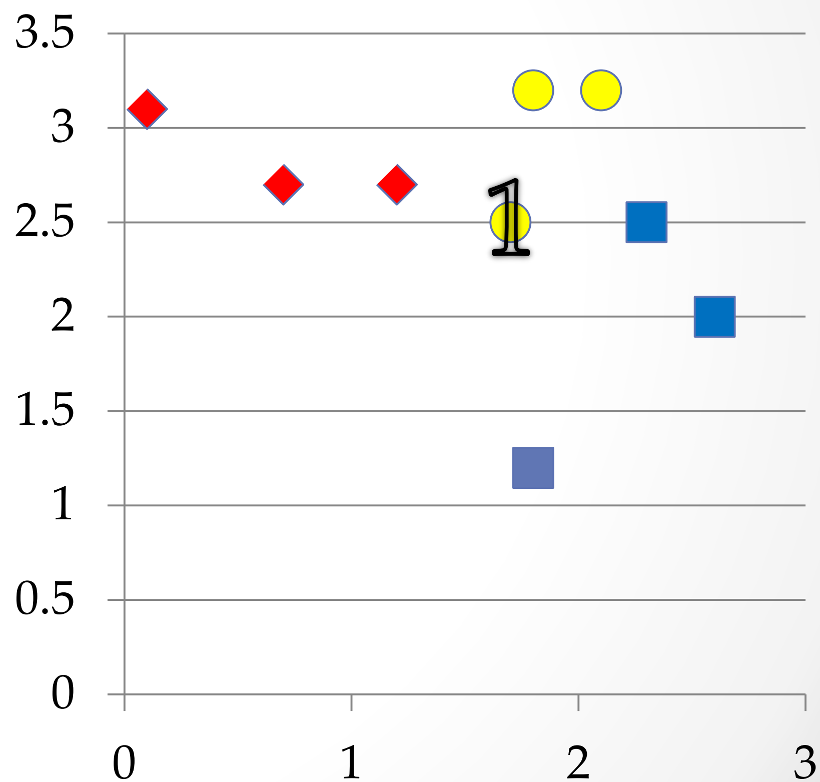
1. User Agent spawns K Clustering Agents ($C = \{c_1, c_2, \cdots, c_K\}$)
2. Each Clustering Agent sends a data request to the indicated Data Agent
3. Data Agent sends first K records ($\{d_1, d_2, \cdots, d_K\}$) to the K Clustering Agents; d_1 to c_1, d_2 to c_2, and so on.
4. Each Clustering Agent calculates its cluster centroid
5. $\forall d_i \in D$ ($i = K + 1$ to n)
6. $\quad \forall c_j \in C$ ($j = 1$ to K)
7. $\quad\quad bidDistance = d_i - \ centroid\ c_j$
8. $\quad\quad$ Allocate d_i to c_j so as to minimise *bidDistance*

Assign Centroid to 1st Cluster Agent

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	?
2.3	2.5	Blue	?
1.8	3.2	Yellow	?
2.6	2.0	Blue	?
0.1	3.1	Red	?
1.2	2.7	Red	?
2.1	3.2	Yellow	?
1.8	1.2	Blue	?

Graphical Depiction

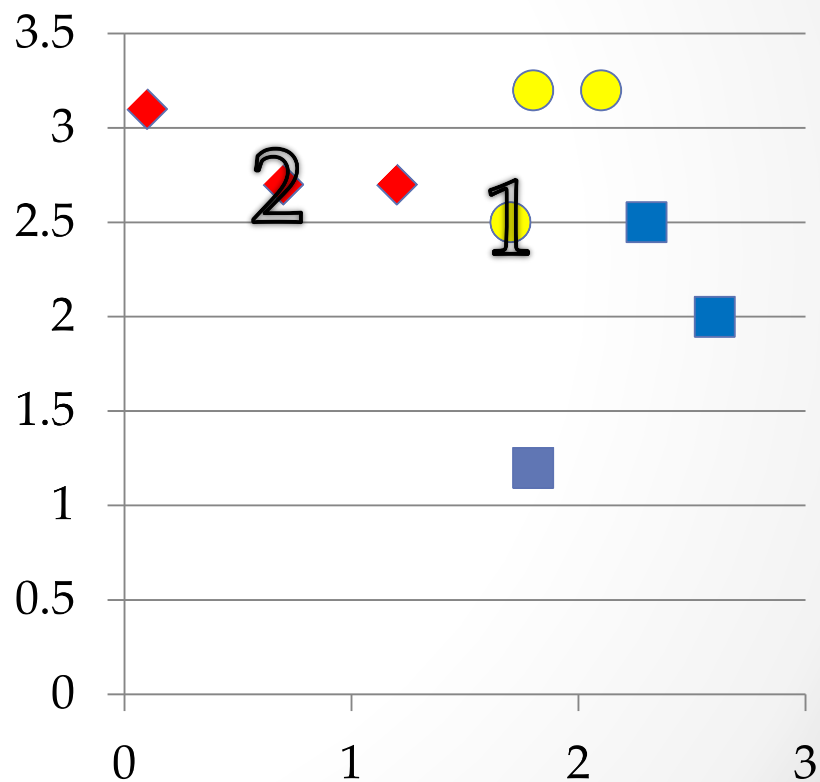


Assign Centroid to 2nd Cluster Agent

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	?
1.8	3.2	Yellow	?
2.6	2.0	Blue	?
0.1	3.1	Red	?
1.2	2.7	Red	?
2.1	3.2	Yellow	?
1.8	1.2	Blue	?

Graphical Depiction

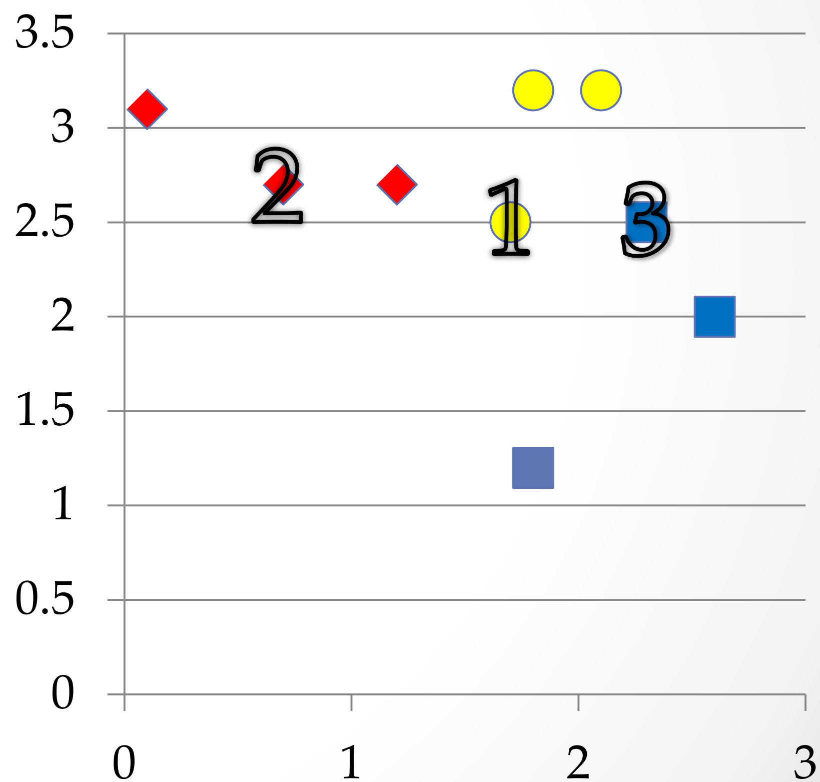


Assign Centroid to 3rd Cluster Agent

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	3
1.8	3.2	Yellow	?
2.6	2.0	Blue	?
0.1	3.1	Red	?
1.2	2.7	Red	?
2.1	3.2	Yellow	?
1.8	1.2	Blue	?

Graphical Depiction

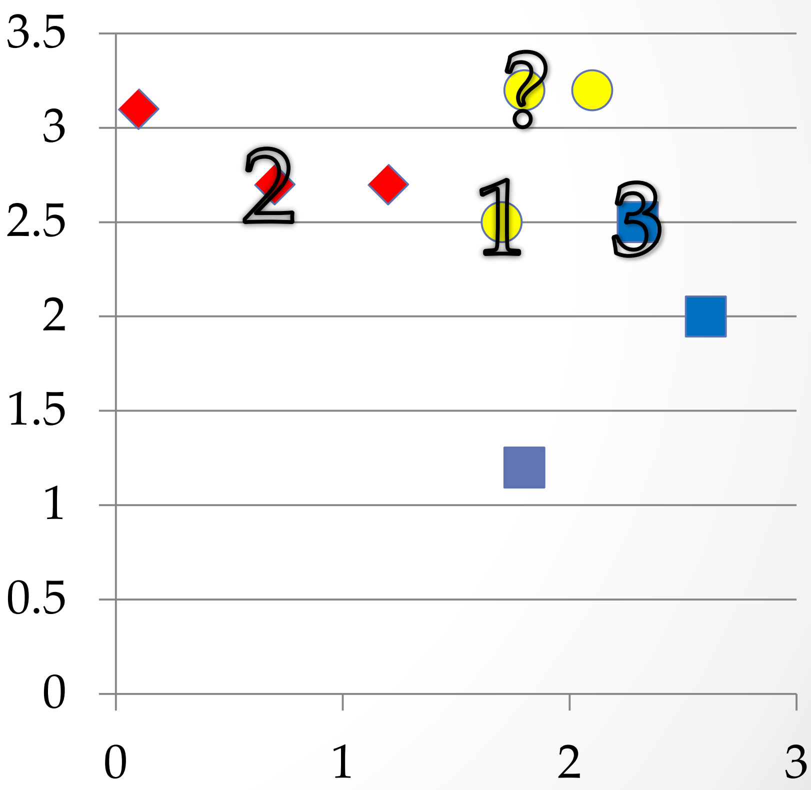


Bid for Instance 4

Bids

Agent	Bid
1	0.707*
2	1.208
3	0.860

Graphical Depiction

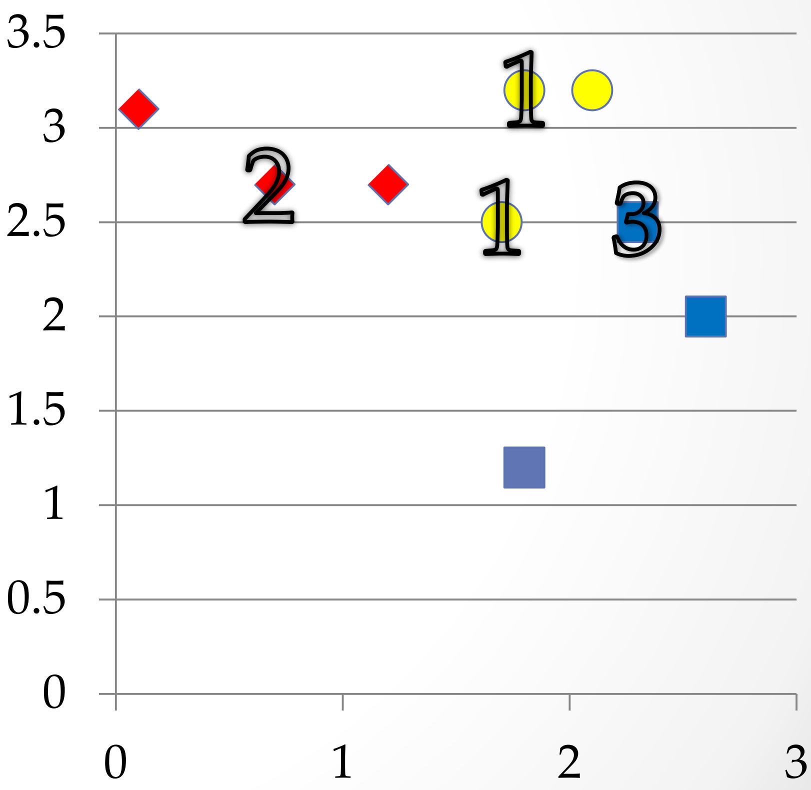


Update

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	3
1.8	3.2	Yellow	1
2.6	2.0	Blue	?
0.1	3.1	Red	?
1.2	2.7	Red	?
2.1	3.2	Yellow	?
1.8	1.2	Blue	?

Graphical Depiction

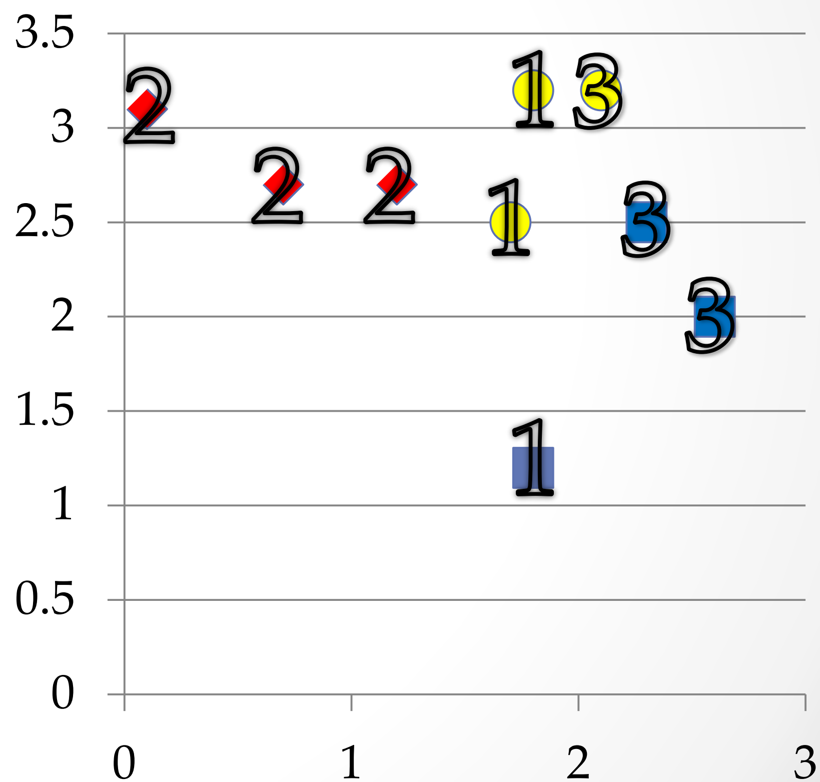


Result of Bidding

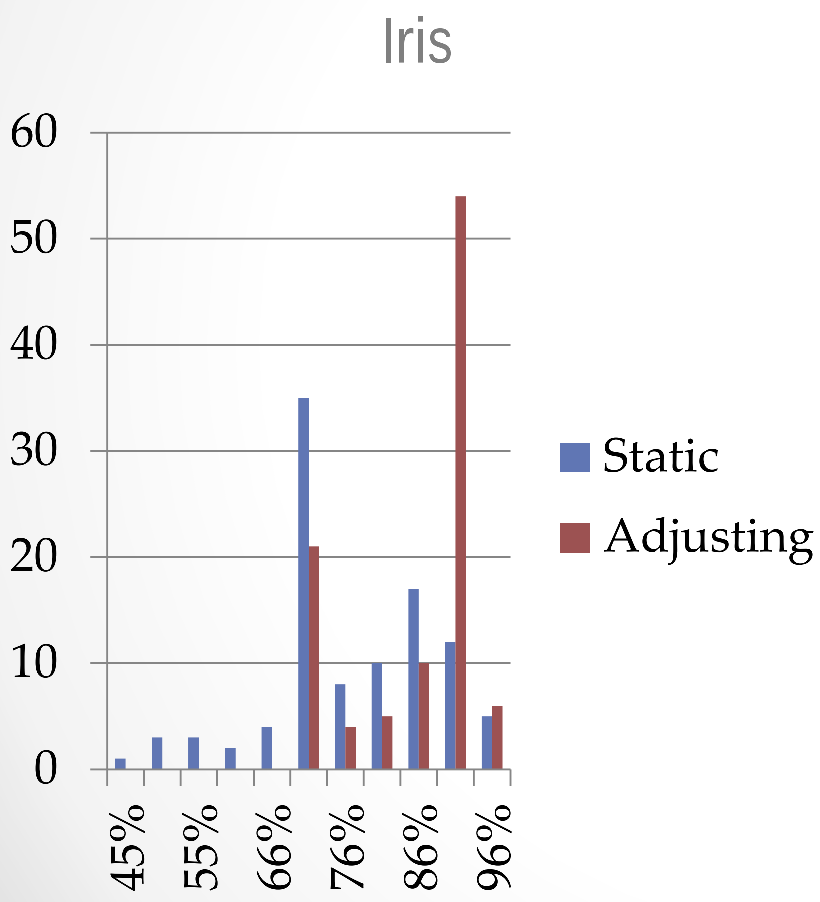
Dataset

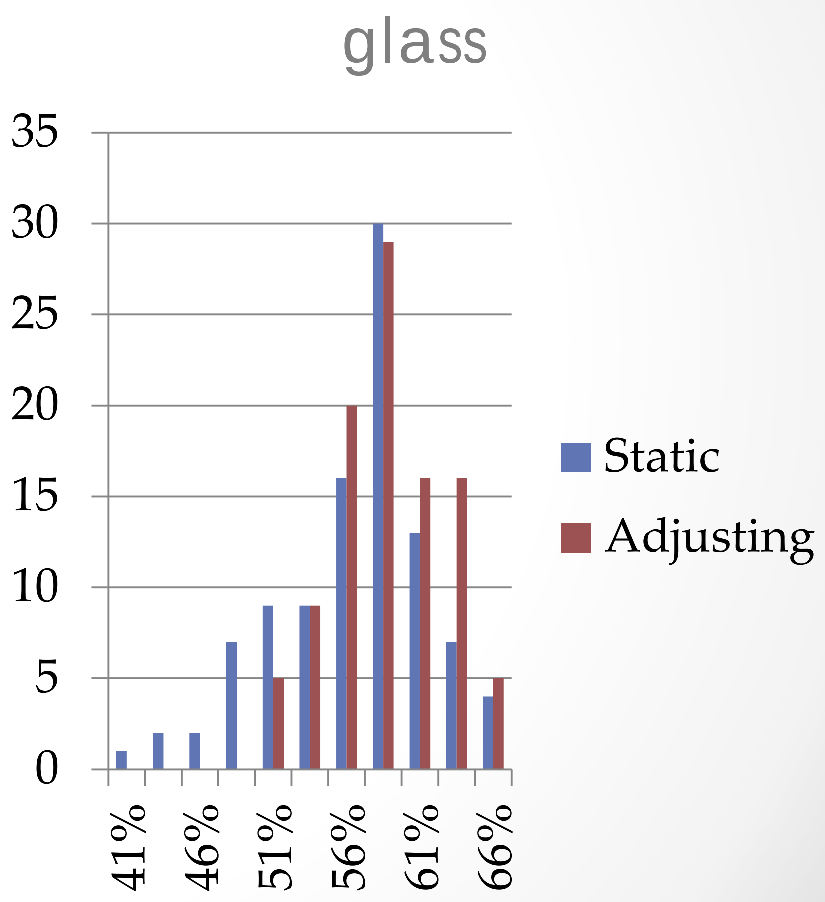
X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	3
1.8	3.2	Yellow	1
2.6	2.0	Blue	3
0.1	3.1	Red	2
1.2	2.7	Red	2
2.1	3.2	Yellow	3
1.8	1.2	Blue	1

Graphical Depiction



Bid Comparison





Refinement Phase

Algorithm Phase II: Refinement

Input: a set of clusters
Output: an improved clustering result

1. For all clusters calculate cluster cohesion and separation values
2. DO WHILE there exists cluster cohesion value > cohesion threshold
   or cluster separation value < separation threshold
3. $\forall c_i \in C$ $(i = 1 \text{ to} K)$
3. Split a cluster c_i into two sub-clusters, c_{major} and c_{minor} using K-means
4. $\forall d \in c_{minor}$
5. $\forall c_j \in C$ $(j = 1 \text{ to } K \text{ and } j \neq i)$
6. $bidDistance = WGAD_j$ (see Section 6)
7. IF $\exists c_j \in C$ such that $bidDistance < cohesion\ threshold$, allocate d
   to c_j so as to minimise $bidDistance$
8. ELSE Allocate d to "outlier cluster"
9. IF no "successful" bids end loop

Gains of Refinement

No.	Data Set	Num Classes	Accuracy Phase I	Accuracy Phase II	Cohesion threshold	Seperation threshold
1	**Iris**	3	0.89	0.97	1.41	2.03
2	**Zoo**	7	0.76	0.78	1.94	1.95
3	**Wine**	3	0.68	0.70	204.95	296.73
4	**Heart**	2	0.55	0.59	33.87	106.28
5	**Ecoli**	8	0.76	0.80	0.42	0.54
6	Blood Transfusion	2	0.76	0.76	794.16	4582.29
7	**Pima Indians**	2	0.65	0.66	65.08	290.60
8	**Breast cancer**	2	0.79	0.85	283.16	1729.93

Calculate Cohesion and Separation

Cohesion

- $WGAD = \frac{\sum_{i=1}^{i=|C|} dist(x_i, c)}{|C|}$

Agent	WGAD
1	0.74
2	0.44
3	0.46
Average	0.54
Target Average	0.44

Separation

- $BGAD = \sum_{i=1}^{i=K} dist(c_i, c)$

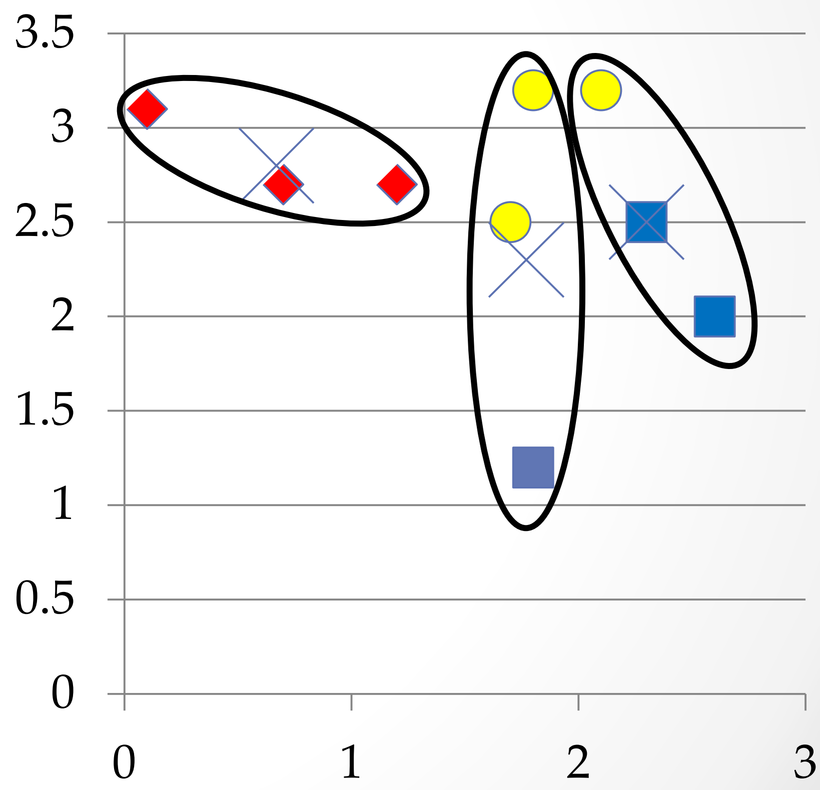
Agent	BGAD
1	1.85
2	2.91
3	2.31
Average	2.36
Target Average	2.83

Recalculate Centroids

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	3
1.8	3.2	Yellow	1
2.6	2.0	Blue	3
0.1	3.1	Red	2
1.2	2.7	Red	2
2.1	3.2	Yellow	3
1.8	1.2	Blue	1

Graphical Depiction

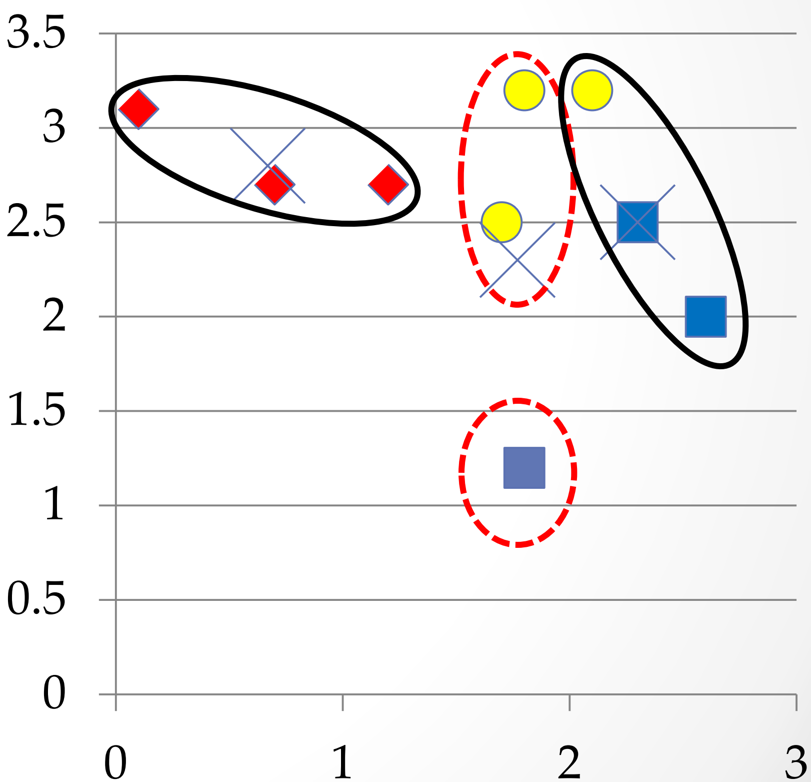


$c_{1_{major}}$ and $c_{1_{minor}}$

WGAD

Agent	$WGAD_j$
2	?
3	?
$cohesion\ threshold = 0.45$	

Graphical Depiction

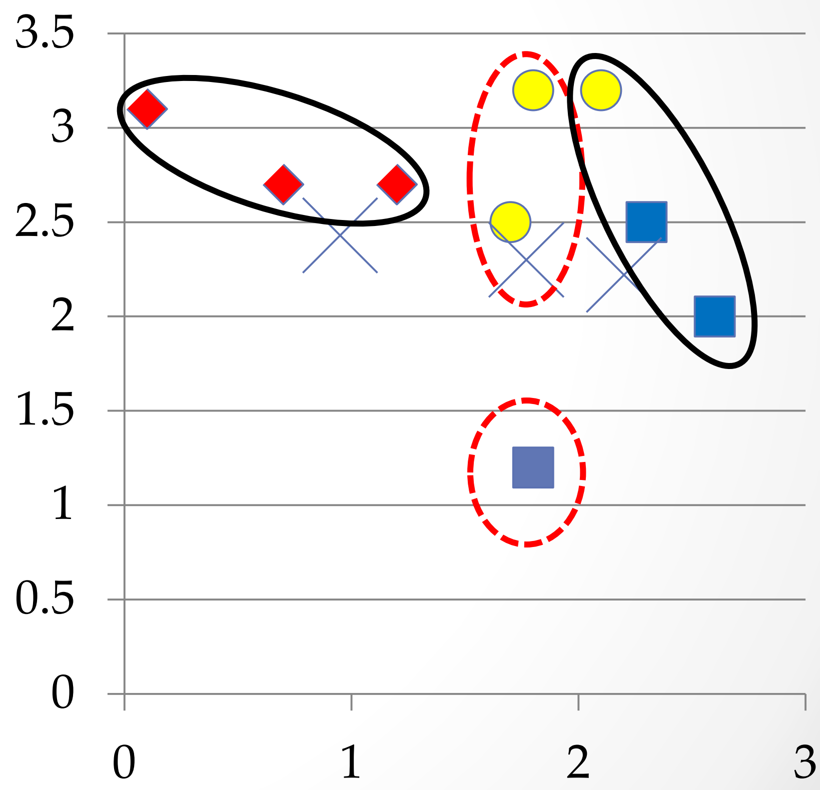


Recalculate Centroids

Centroids

Agent	X-axis	Y-axis	$WGAD_j$
2	0.95	2.43	0.83
3	2.2	2.22	0.71
$cohesion\ threshold = 0.45$			

Graphical Depiction

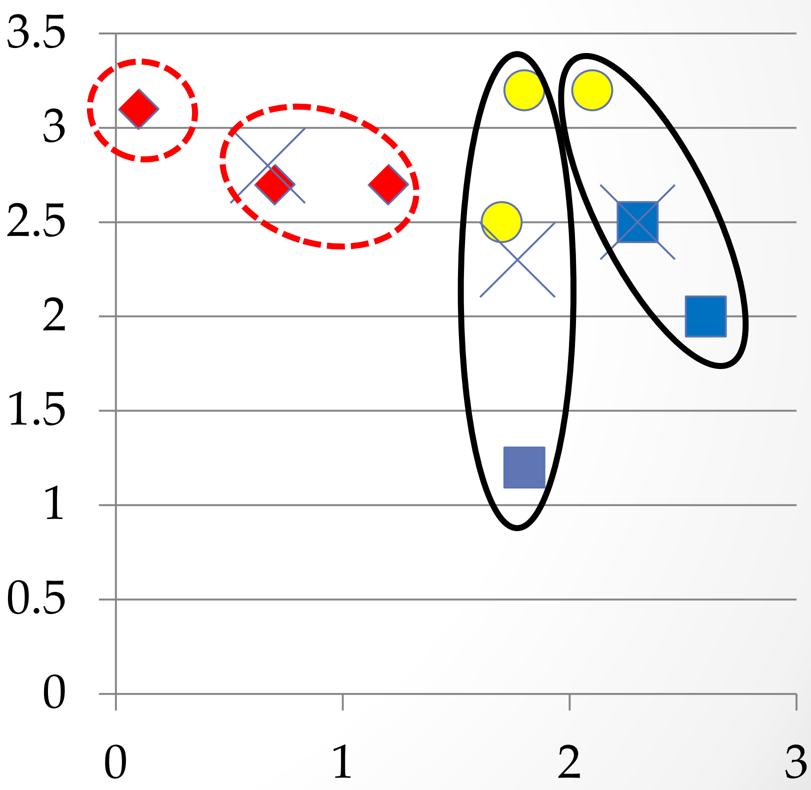


$c_{2_{major}}$ and $c_{2_{minor}}$

$WGAD_{\mathrm{j}}$

- Obvious that $WGAD_j$ will not be minimized for Agent 1 nor Agent 3

Graphical Depiction

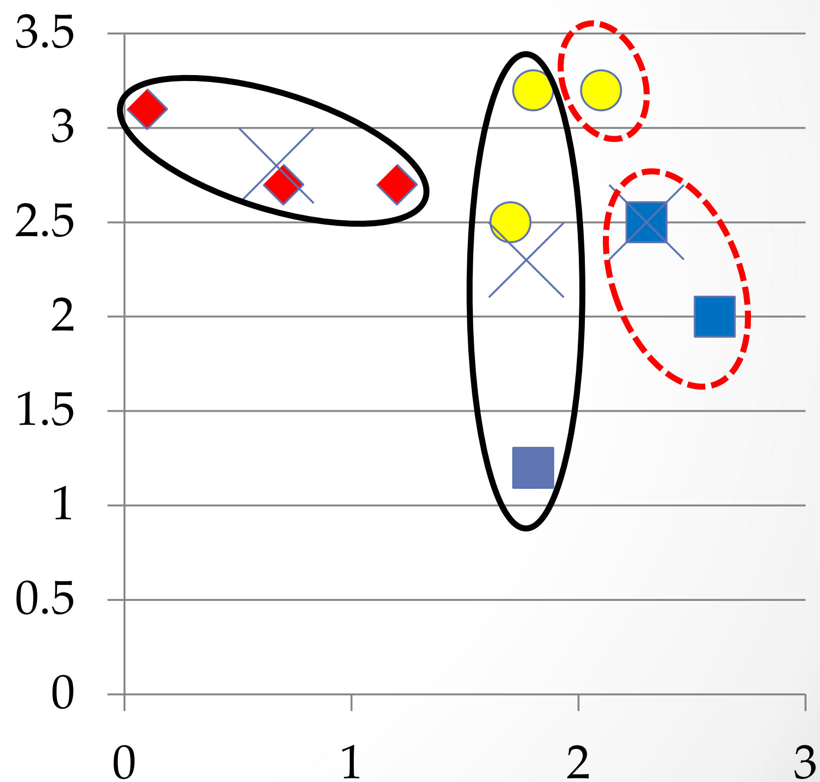


$c_{3_{major}}$ and $c_{3_{minor}}$

Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	3
1.8	3.2	Yellow	1
2.6	2.0	Blue	3
0.1	3.1	Red	2
1.2	2.7	Red	2
2.1	3.2	Yellow	3
1.8	1.2	Blue	1

Graphical Depiction

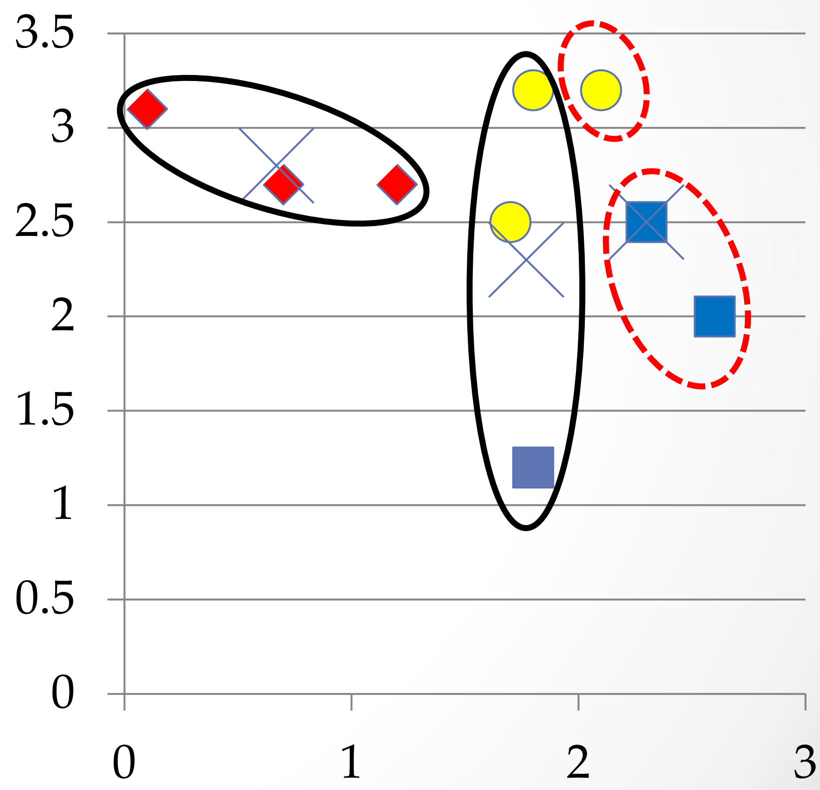


Recalculate Centroids

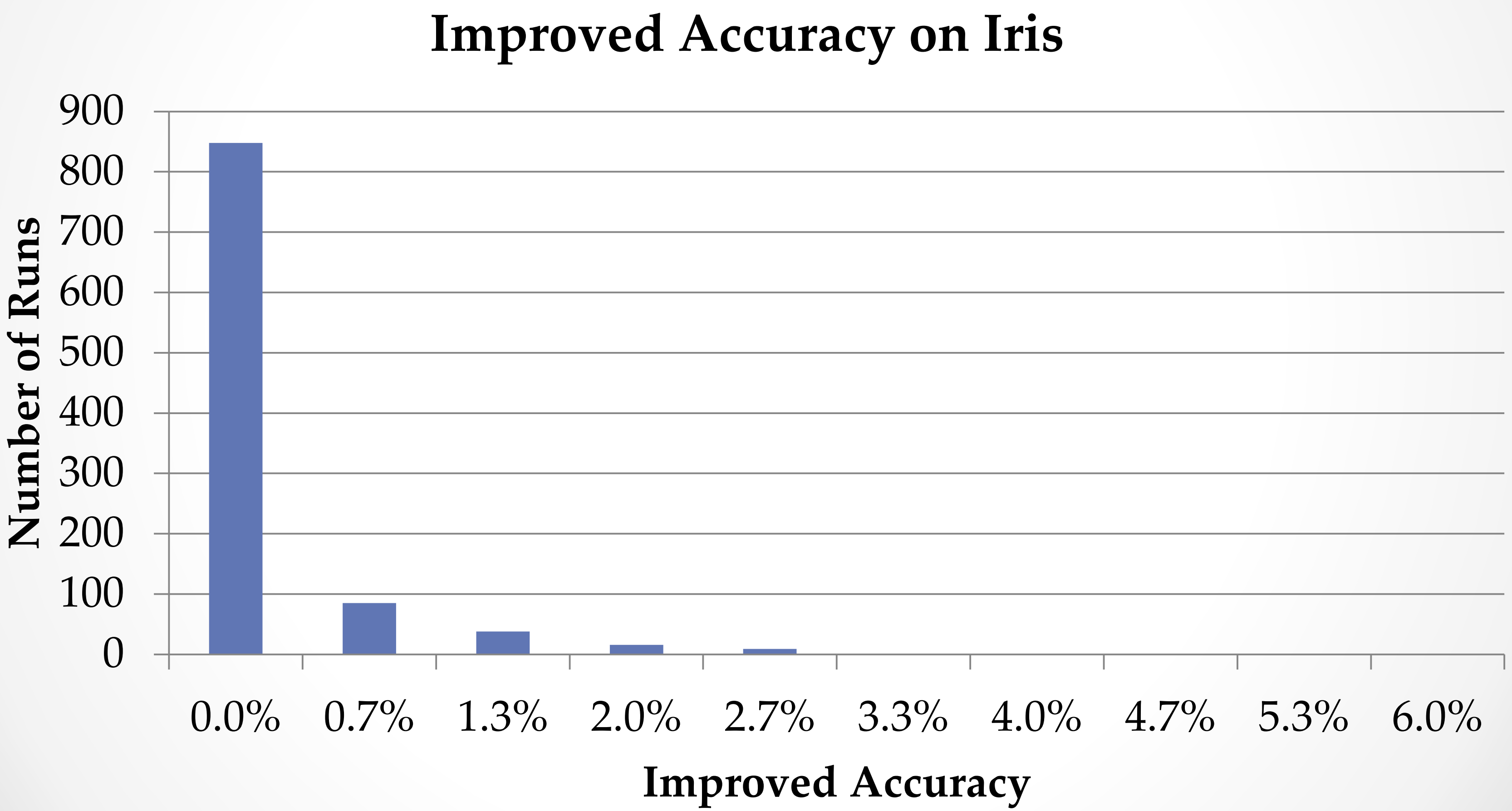
Centroids

Agent	X-axis	Y-axis	$WGAD_j$
1	1.85	2.53	0.72
2	1.03	2.93	0.74
$cohesion\ threshold = 0.45$			

Graphical Depiction



"Improvement" of Refinement Phase



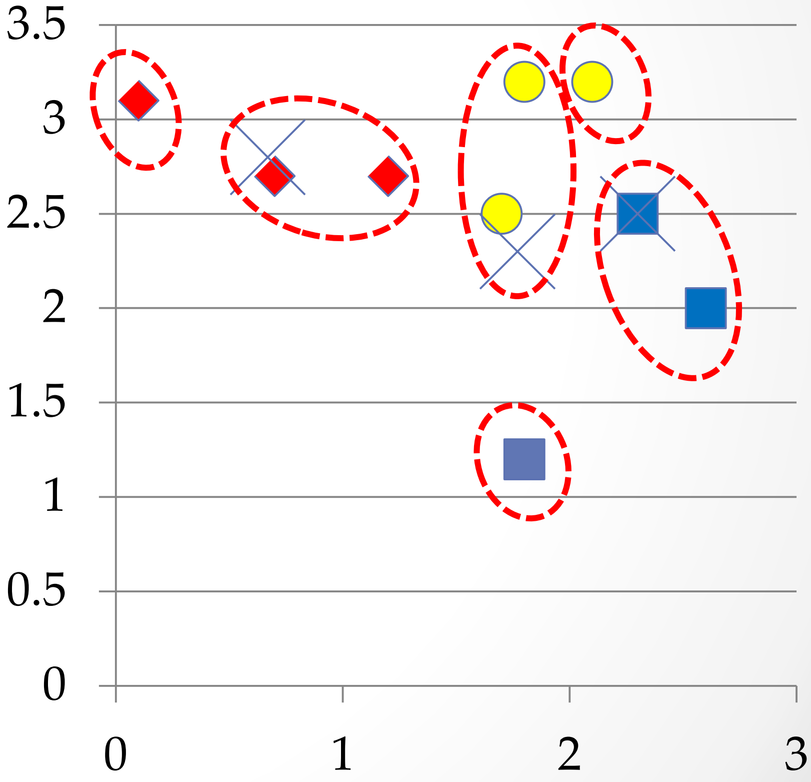
quasi

Simultaneous Bidding

Up for Bid

Agent	X-axis	Y-axis
1	1.8	1.2
2	0.1	3.1
3	2.1	3.2

Graphical Depiction

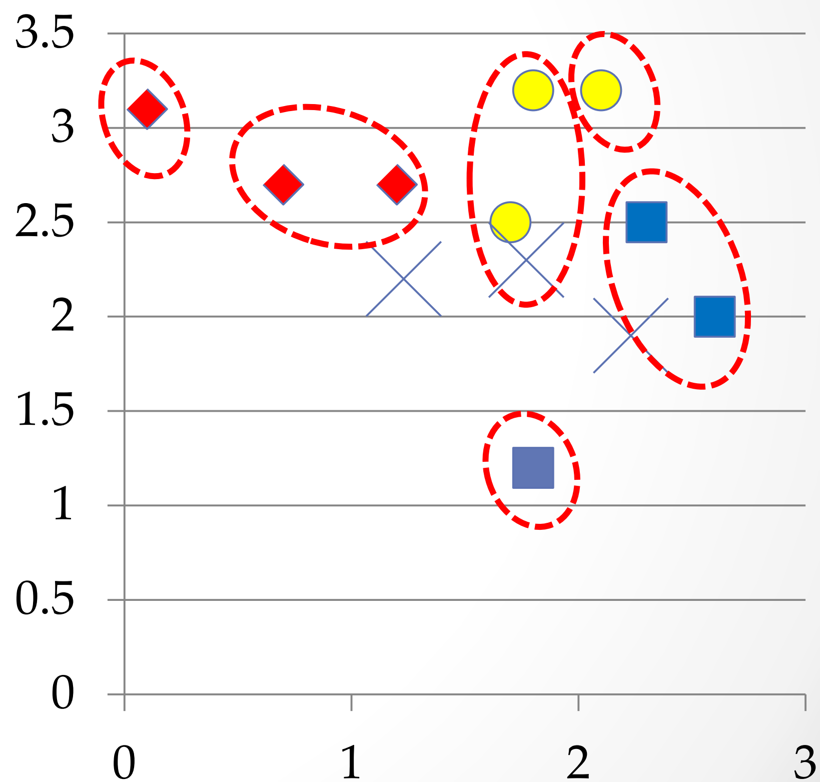


Bidding on (1.8, 1.2)

Recalculate Centroids

Agent	X-axis	Y-axis	$WGAD_j$
2	1.23	2.2	0.79
3	2.23	1.9	0.60
$cohesion\ threshold = 0.45$			

Graphical Depiction

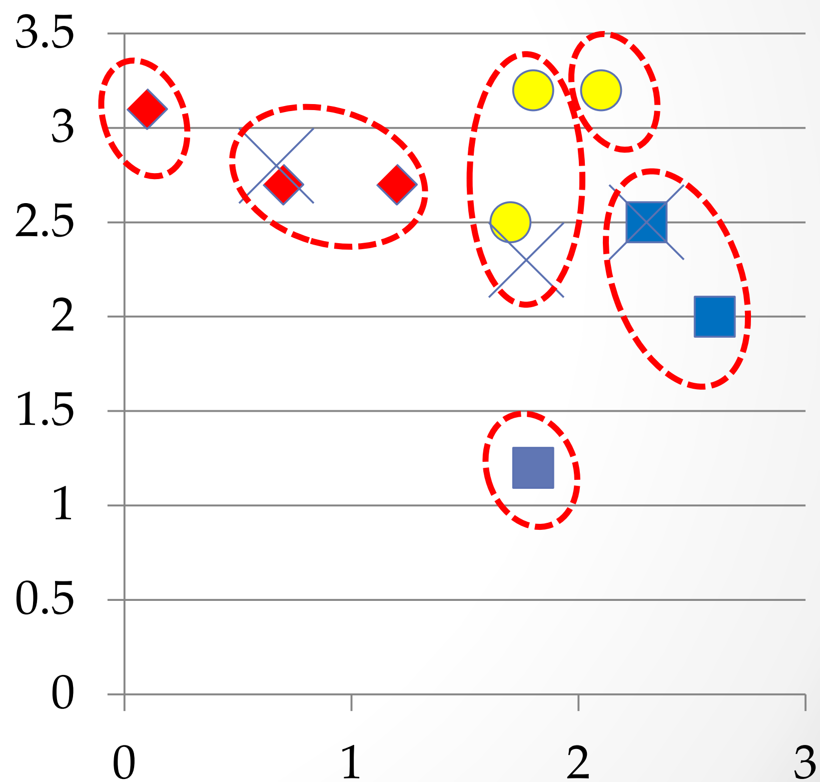


Pass on (0.1, 3.1)

Justification

- Obviously will not help Agents 1 or 3
- Agent 2 cannot bid on the $c_{2_{minor}}$

Graphical Depiction

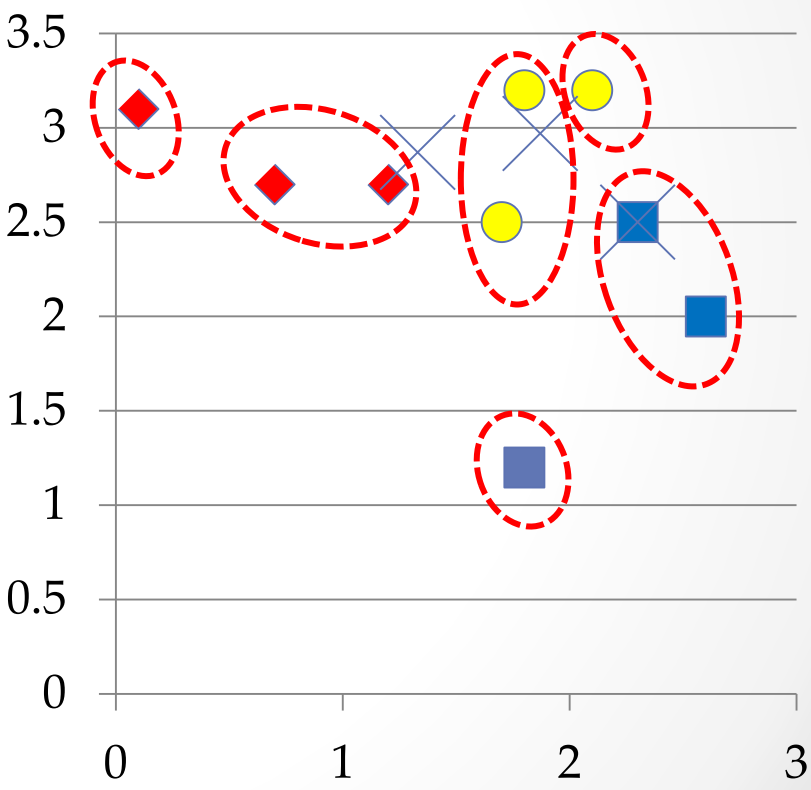


Bidding on (2.1, 3.2)

Recalculate Centroids

Agent	X-axis	Y-axis	$WGAD_j$
1	1.87	2.97	0.36*
2	1.33	2.87	0.57
$cohesion\ threshold = 0.45$			

Graphical Depiction

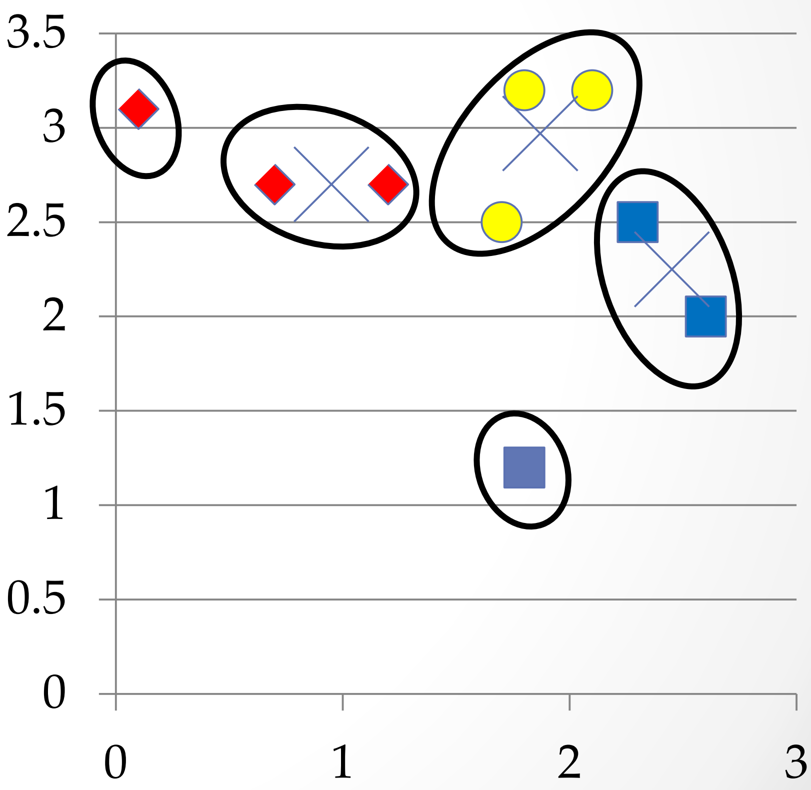


Clusters after 1 Iteration

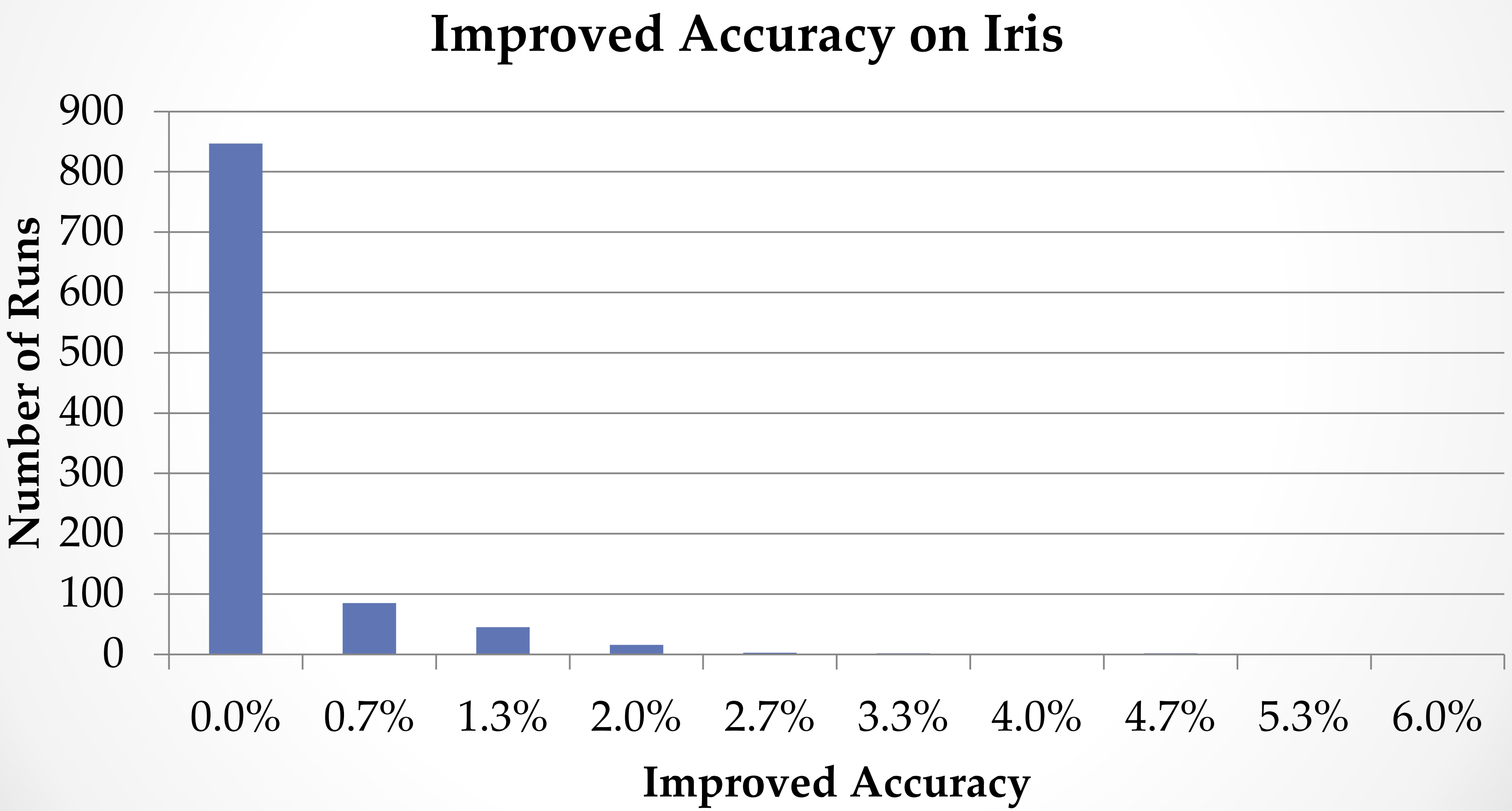
Dataset

X-axis	Y-axis	Color	Agent
1.7	2.5	Yellow	1
0.7	2.7	Red	2
2.3	2.5	Blue	3
1.8	3.2	Yellow	1
2.6	2.0	Blue	3
0.1	3.1	Red	Outlier
1.2	2.7	Red	2
2.1	3.2	Yellow	1
1.8	1.2	Blue	Outlier

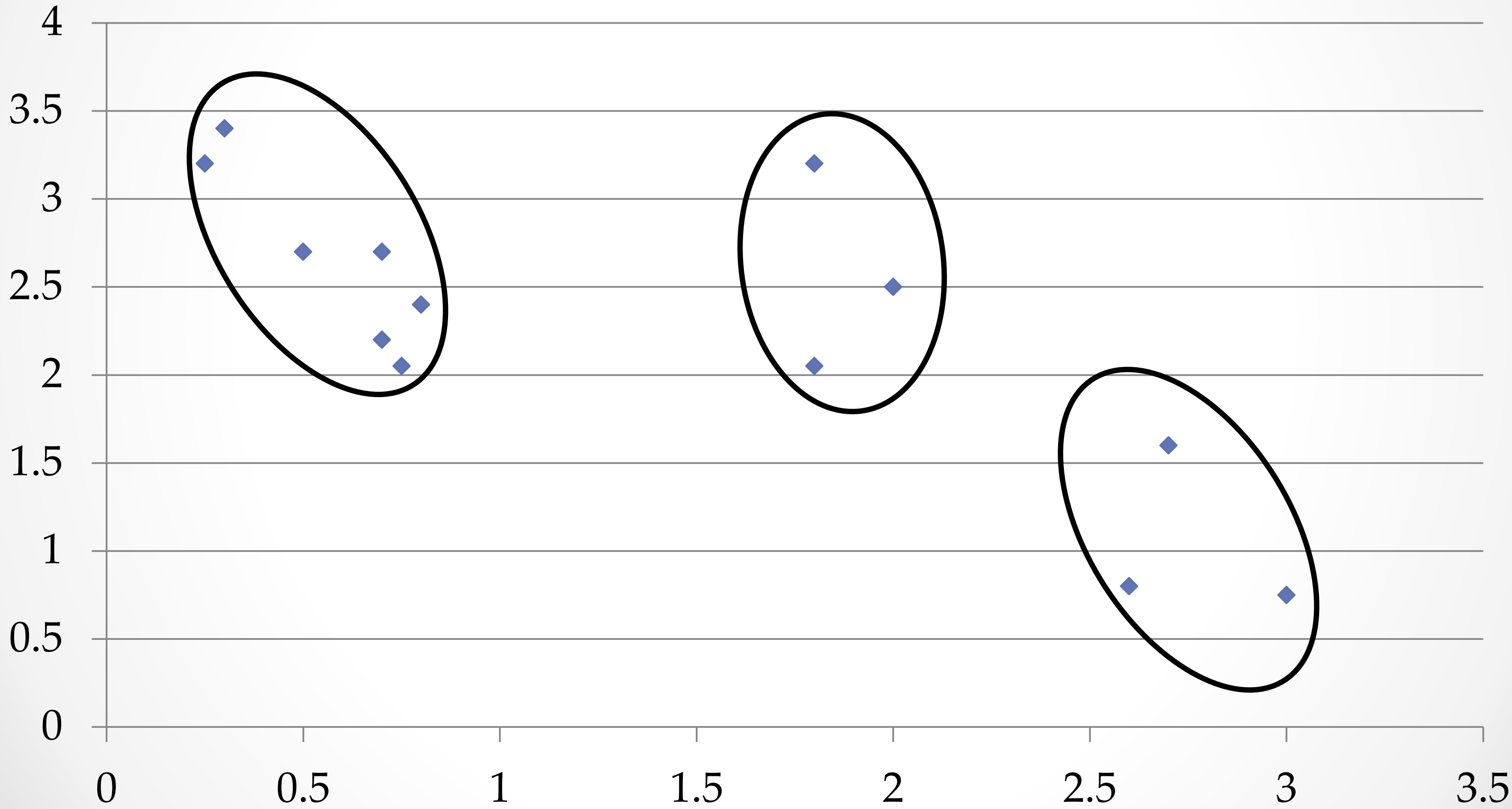
Graphical Depiction



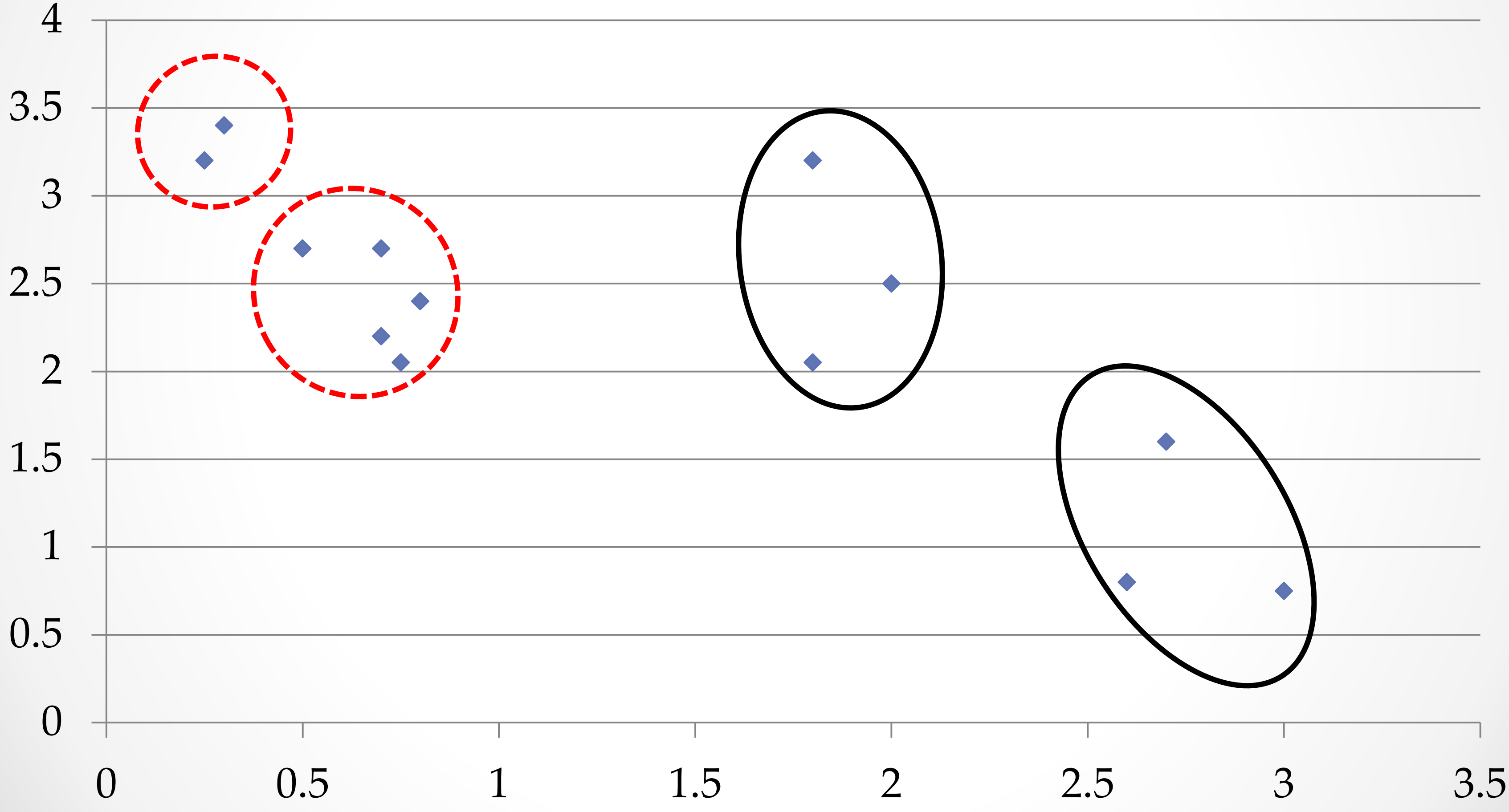
Improvement?



Justification for *k*Means in inner loop?



Justification for *k*Means in inner loop?



Instead of returning c_{major} and c_{minor}…

Method #1

- For each cluster:
  - Return the instance that is furthest from centroid… with some probability
  - Other Clusters then bid on this instance based on whether included the instance increases cluster $wgad$
  - Return to original cluster otherwise

Method #2

- For each cluster c_i:
  - Return k instances that are furthest from the centroid
  - Bid on the instances in the same manner as in Method #1

Benchmarks

Algorithm	Iris	Glass	Vehicle
MLP	97.33%	67.76%	81.80%
J48	96.00%	66.82%	72.58%
Simple *k*-Means	88.67%	44.86%	36.41%
ZeroR	33.33%	35.51%	25.65%

Parameters

Method #1

- Number of auctions
  - [1,∞]
    - Choose 20
- Probability of considering a further point
  - [0,1]
    - {.1, .2, .3, .4}

Method #2

- Number of auctions
  - [1,1000]
    - Choose 20
- Number of instances in c_{minor}
  - {2,3,4,5}

Results: Method 1

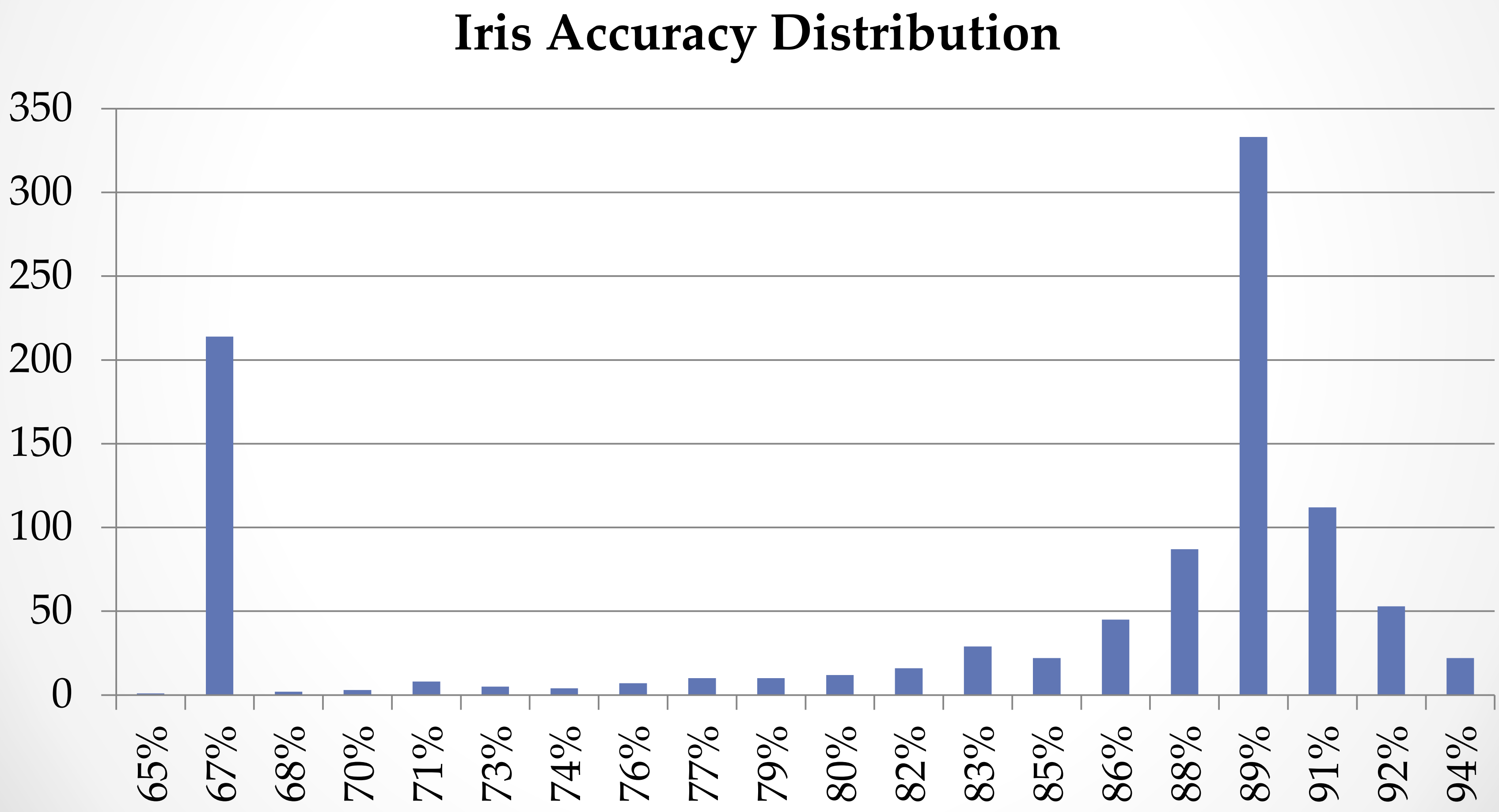
Accuracy

Data	0.1	0.2	0.3	0.4
Iris	83.1%	83.7%	81.9%	82.8%
Glass	57.6%	57.0%	57.6%	56.9%
Vehicle	43.6%	43.4%	43.3%	43.3%

Standard Deviation

	0.1	0.2	0.3	0.4
Iris	9.08%	9.06%	9.57%	9.17%
Glass	4.23%	4.14%	4.23%	4.24%
Vehicle	2.40%	2.16%	2.31%	2.37%

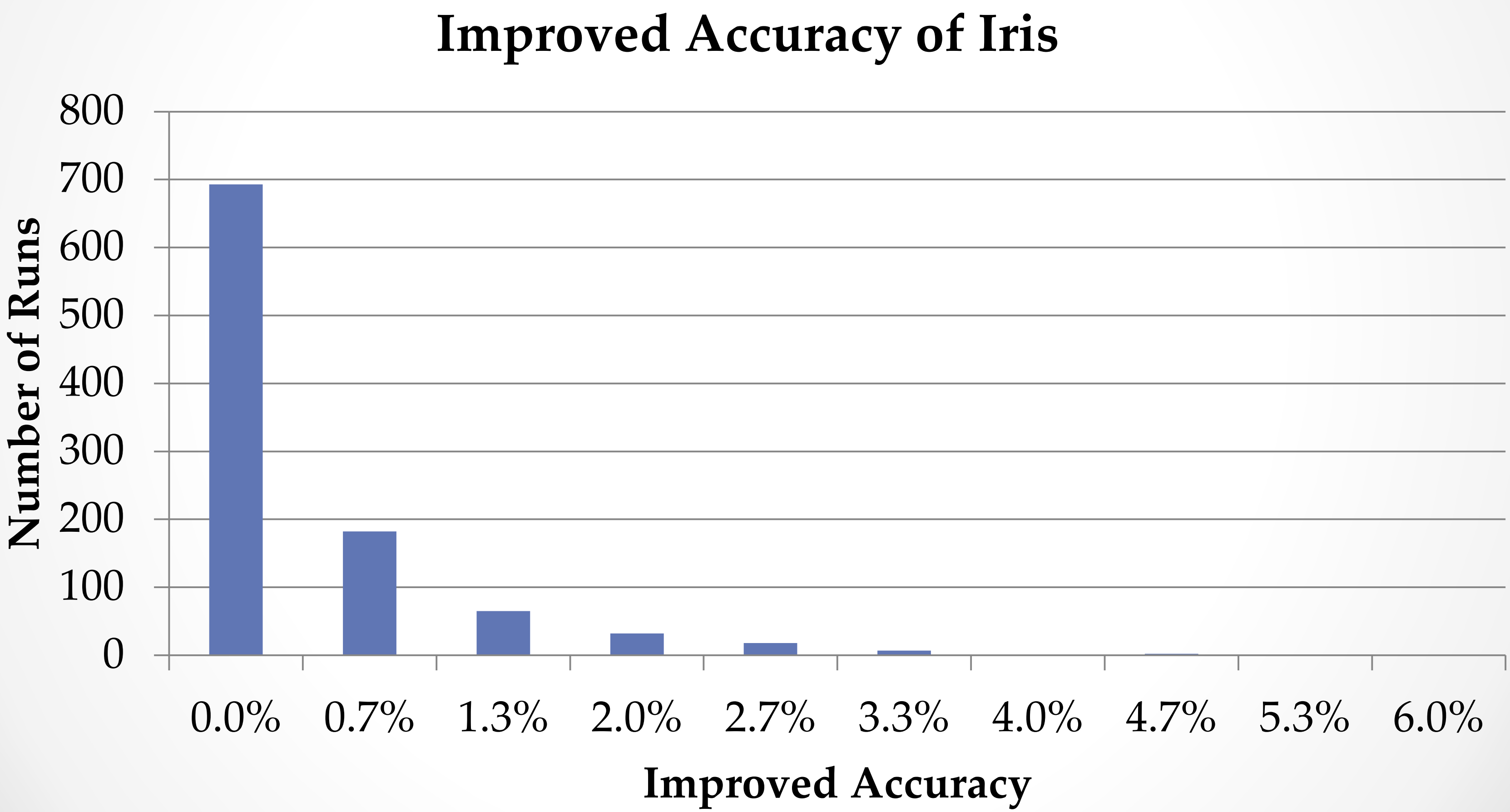
Highly Dependent on Bid



Refinement lacks Refinement

Dataset	After Bid Process	After Refinement
Iris	83.43%	83.75%
Glass	57.61%	57.04%
Vehicle	43.36%	43.40%

Small Improvement in Refinement



Results Method #2

Dataset%	Accuracy	Standard Deviation
Iris	82.94%	9.27%
Glass	57.72%	3.57%
Vehicle	43.47%	2.31%

Mining Sequence Data and Time Series Data

- Information/search agents to get information
    - “assimilation of information into inputs to predictor agents.”
- Learner/miner to modify information selection criteria
    - “apportioning of biases to feedback”
    - “developing rules for Search Agents to collect information”
    - “developing rules for Information Agents to assimilate information.”
- Predictor agents to predict trends
    - “Incorporation of qualitative information”
    - “Multi-objective optimization not in closed form”

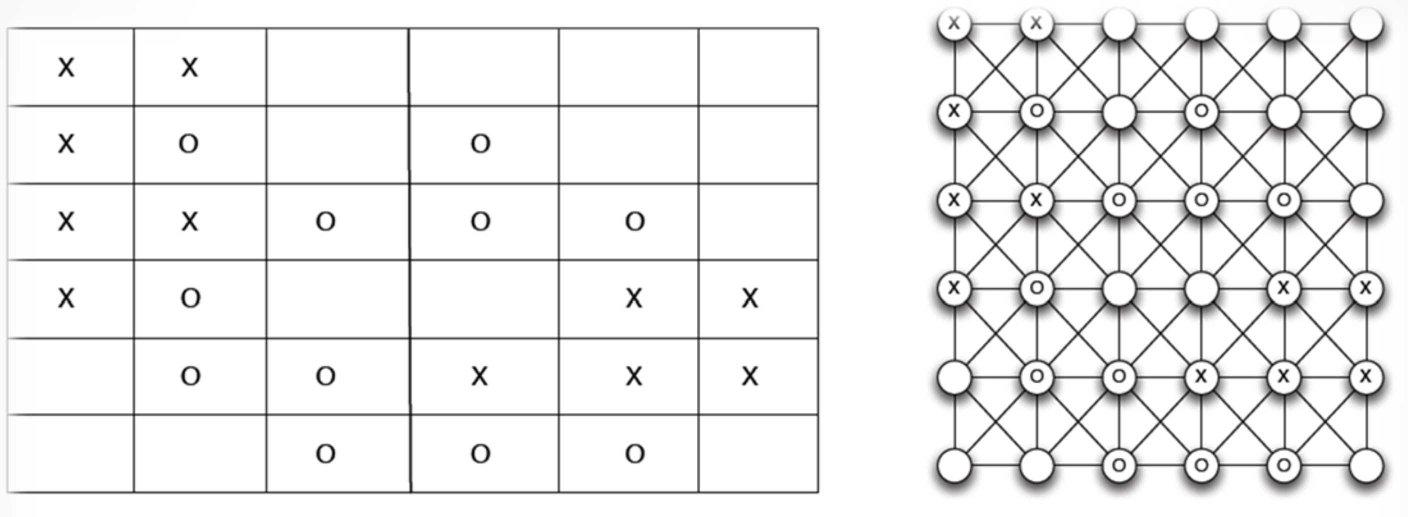
Mining Complex Knowledge from Complex Data

- Complexity from data that are non-i.i.d (independent and identically distributed)
    - “In most domains, the objects are not independent of each other, and are not of a single type.”
    - “We need data mining systems that can soundly mine the rich structure of relations among objects”
        - Web pages
        - Social networks
        - Metabolic networks in cells
- Recognize movements of objects and people from Web and data logs
    - “[Find] useful spatial and temporal knowledge.”
- Biggest performance gap for data mining systems
    - “[inability] to relate the result of mining to the real-world decision they affect-all they can do is hand the results back to the user.”

Data Mining in a Network Setting

- Community mining and mining of social networks has become a key area of research
- Identifying community structures
    - “it’s critical to have the right characterization of the notion of ‘community’ that is to be detected”
    - “the entities/nodes involved are distributed in real-life applications, and hence distributed means of identification will be desired.”
    - “a snapshot-based dataset may not be able to capture the real picture
        - “what is most important lies in the local relationships between entities/nodes.”
- 2 Challenges. Modeling;
    - “the network’s static structures (e.g. topologies and clusters)”
    - “dynamic behavior (such as growth factors, robustness, and functional efficiency)”

Consider the Following MAS



(a) *Agents occupying cells on a grid.* (b) *Neighbor relations as a graph.*

Initial Configuration

X1*	X2*				
X3	O1*		O2		
X4	X5	O3	O4	O5*	
X6*	O6			X7	X8
	O7	O8	X9*	X10	X11
		O9	O10	O11*	

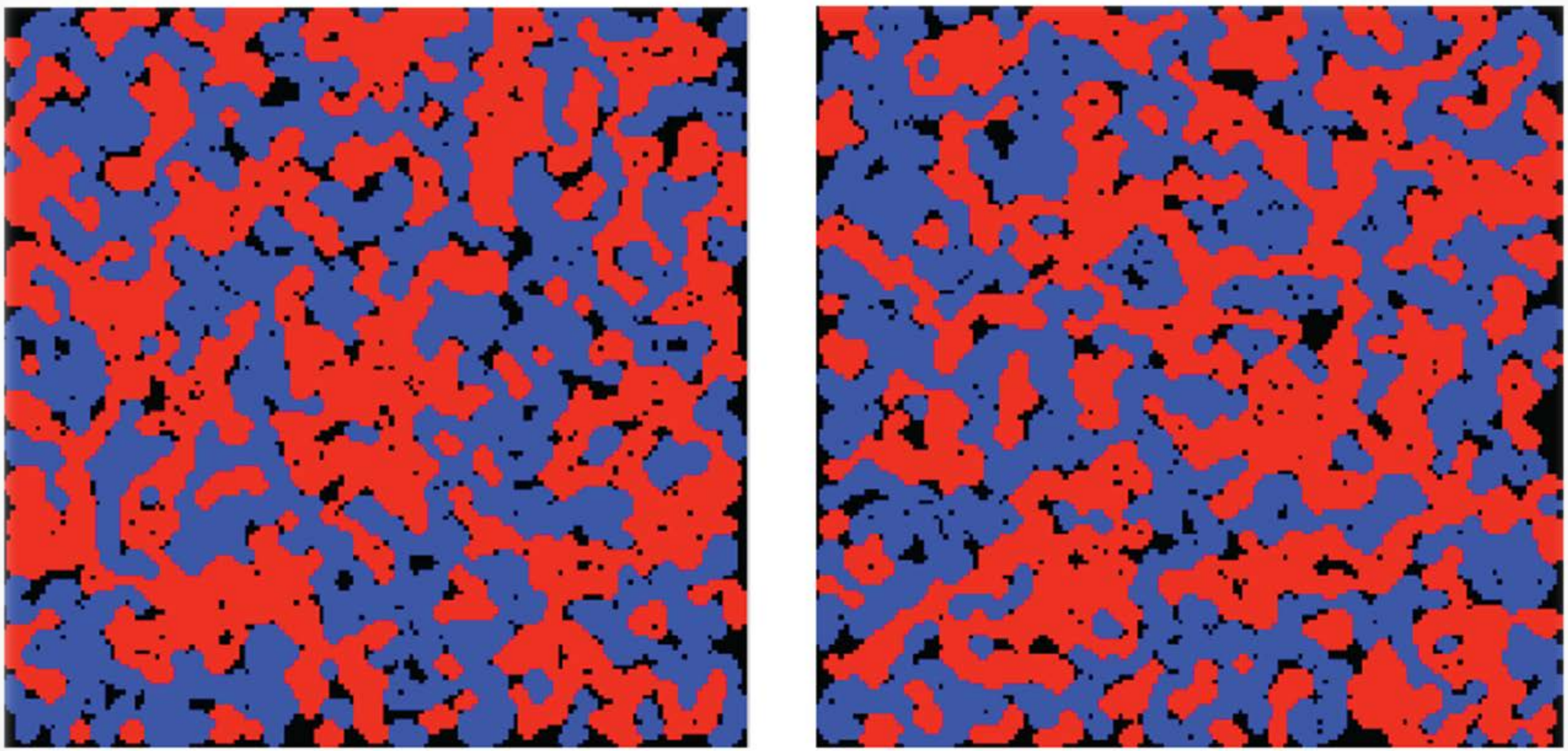
(a) *An initial configuration.*

After One Round

X3	X6	O1	O2		
X4	X5	O3	O4		
	O6	X2	X1	X7	X8
O11	O7	O8	X9	X10	X11
	O5	O9	O10*		

(b) *After one round of movement.*

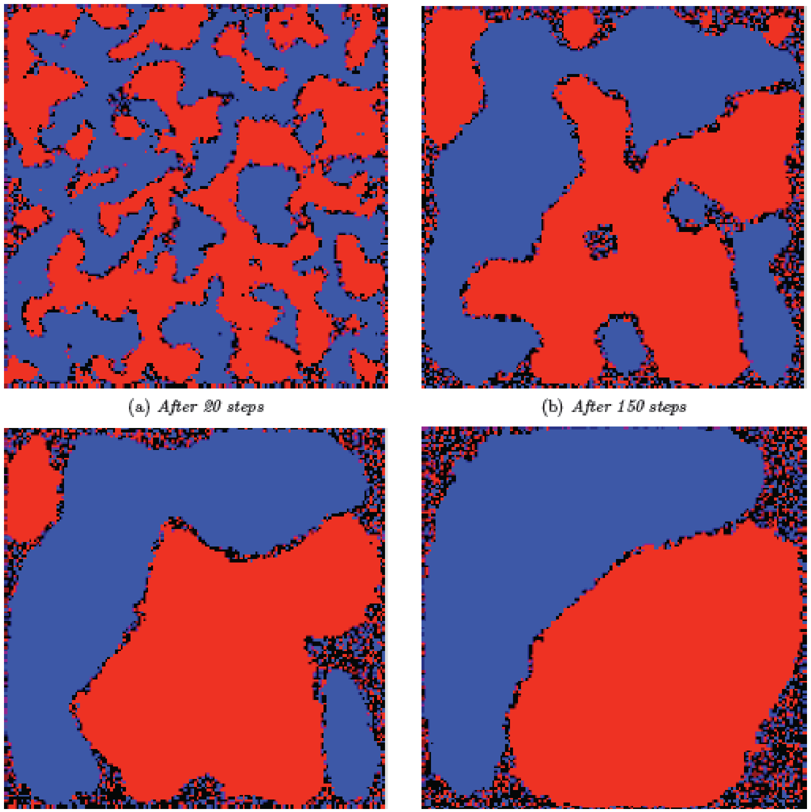
Threshold = 3



(a) *A simulation with threshold 3.*

(b) *Another simulation with threshold 3.*

Threshold = 4



(a) *After 20 steps* (b) *After 150 steps*

(c) *After 350 steps* (d) *After 800 steps*

Distributed Data Mining and Mining Multi-Agent Data

- Minimize quantity of data shipped between various sites
- Adversaries that deliberately manipulate data
    - o "We need to develop systems that explicitly take this into account, by combining data mining with game theory."

Scale Down Data

- Distributed data
  - Select relevant data from each location
  - Move only the local patterns
- Instance selection
  - Simultaneously
    - Reduce size of data
    - Preserve extractable information
    - Produce a comparable classifier
- Taxonomies
  - Filter, wrapper, embedded methods
  - Incremental search, decremental search, batch search

Agent-Based Approach

Algorithm 3.2 Local search for instance selection.

Input: s-individual representing a solution encoded as a string consisting of numbers of selected reference instances; L-list of the instance numbers not in s; t-number of clusters of potential reference instances in s.
Output: $solution$-the improved individual.

1. Set i by drawing it at random from $\{1, 2, \ldots, t\}$.
2. Identify j which is an instance number representing the ith cluster.
3. Set j' by drawing it at random from L.
4. Replace an instance numbered j by an instance numbered j' within the ith cluster of s thus producing individual s'.
5. Calculate fitness of s'.
6. If (s' is better then s) then ($s := s'$ AND j replaces j' in L).
7. If ($!terminating\_condition$) then go to 1.
8. $solution := s$.

Data Mining for Biological and Environmental Problems

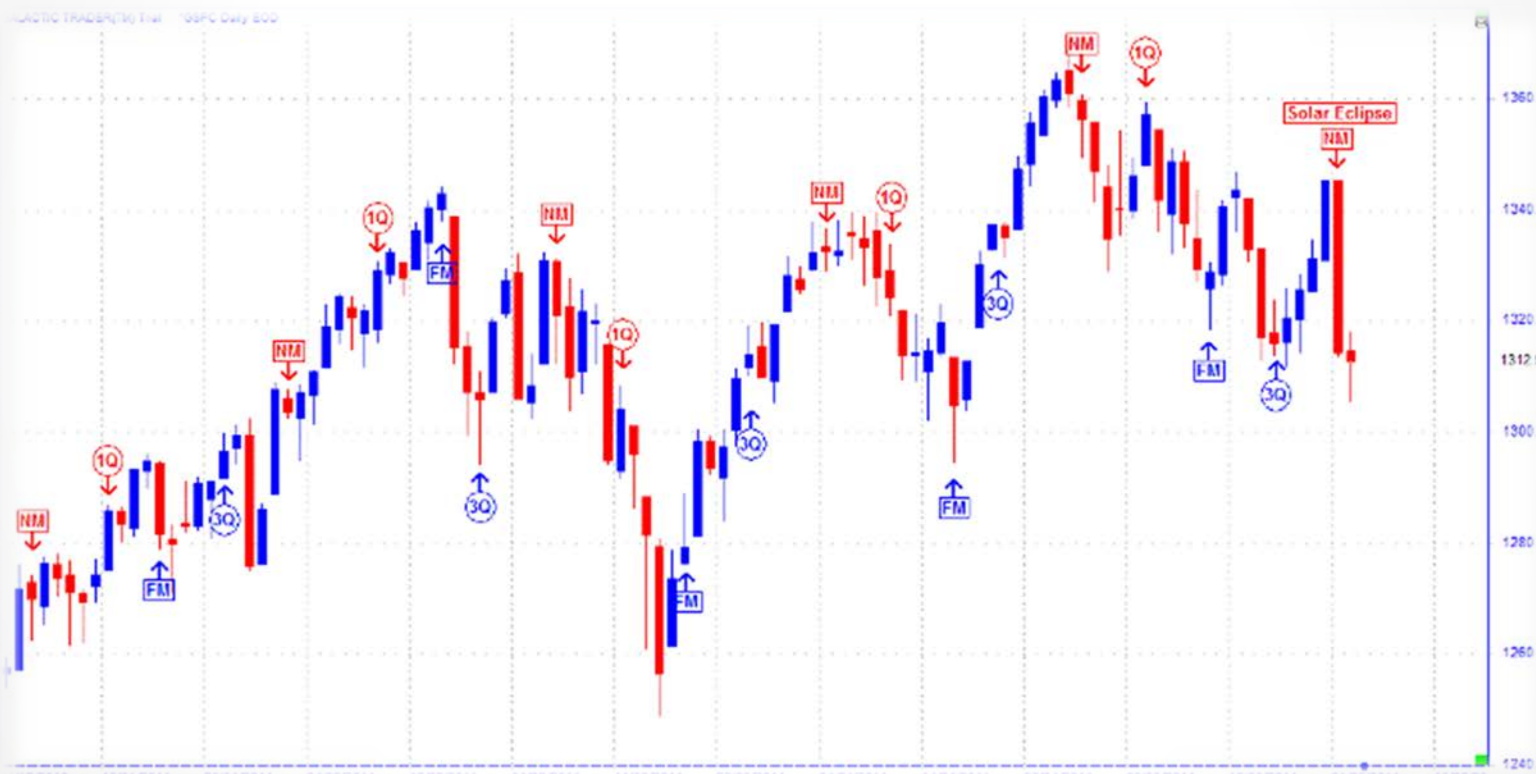
- "In molecular biology, many complex data mining tasks exist, which cannot be handled by standard data mining algorithms."
  - DNA
  - Chemical properties
  - 3D structures
  - Functional properties
- One the great challenges is dealing with "dynamic temporal behavioral pattern identification and prediction in"
  - Very large scale systems (global climate changes, "bird flu", etc…)
  - Human-centered systems (User-adapted human-computer interaction, P2P, etc…)

All Systems seek equilibrium

- "In the natural world, everything seeks stability, which means seeking a state of minimal energy."
- "[J]ust as in a chemical reaction all atoms are simultaneously seeking a state of with minimum energy, in an economy all people are seeking to maximize their utility."
- "And now, my beloved son, notwithstanding their hardness, let us *labor diligently*; for if we should cease to labor, we should be brought under condemnation; for we have a *labor* to perform whilst in this tabernacle of clay, that we may *conquer* the enemy of all righteousness, and *rest* our souls in the kingdom of God." Moroni 9:6

Leafcutter Ants

Simulating long memory in the stock market



Rayner et al

General Electric

www.google.com/finance

LinkedIn Corp

www.google.com/finance

Long memory

- Over periods of time volume can be consistently high or low
- Similar volatilities appear in the market in clusters
- Buying/selling like volatility and volume exhibit long memory
- Returns do not exhibit long memory, similar returns do not cluster together, and high frequency returns exhibit anti-persistence

3 Strategies

- Fundamentalists – value stock based on perceived long-term value.
- Chartists – valuate based on historical data.
- Noise traders – forecasts based on what is believed to be a noise signal.
    - "a stock trader whose decisions to buy, sell, or hold are irrational and erratic. The presence of noise traders in financial markets can then cause prices and risk levels to diverge from expected levels even if all other traders are rational." (https://en.wikipedia.org/wiki/Noise_trader)

Agent components

- $\hat{p}_{(i,t+\tau)} = p_t \cdot e^{\hat{r}(i,t,t+\tau)}$
  - Where p_t is quoted price at time t and agent will *buy* iff $\hat{p}_{(i,t+\tau)} \geqslant p_t$ and *sell* otherwise
- $\hat{r}_{(i,t,t+\tau)} = \hat{r}_{f(i,t,t+\tau)} + \hat{r}_{c(i,t,t+\tau)} + \hat{r}_{n(i,t,t+\tau)}$
- where
  - $\hat{r}_{f(i,t,t+\tau)} = f_{(i,t)} \cdot \left(\frac{F-p_t}{p_t}\right)$
  - $\hat{r}_{c(i,t,t+\tau)} = c(i,t) \cdot \hat{r}_{L_i}$
  - $\hat{r}_{n(i,t,t+\tau)} = n(i,t) \cdot \epsilon(i,t)$
- Note that p_t is price at previous time step, $\epsilon(i,t)$ is random variable distributed $\sim N(0,1)$, and r_{L_i} is forecast based off historical data

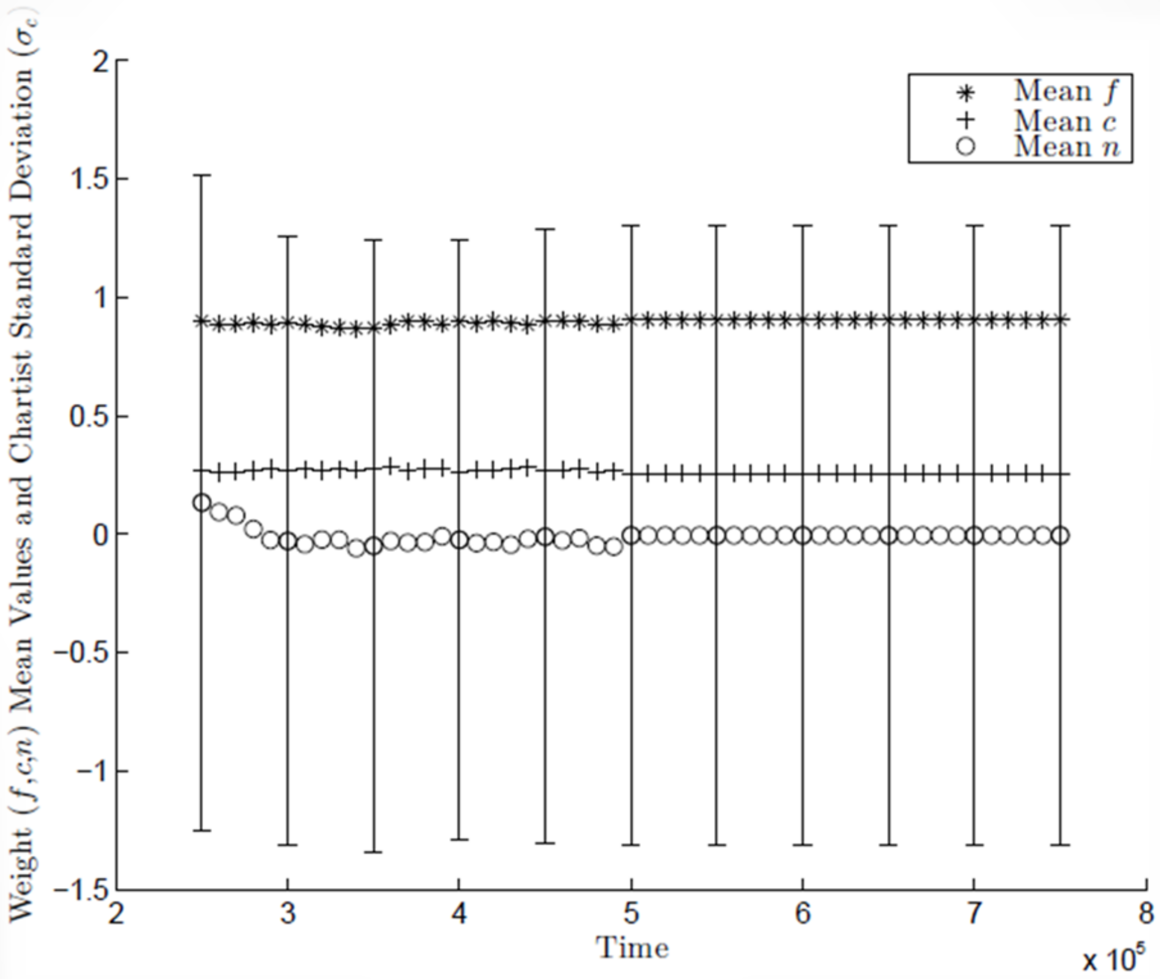
Co-evolution

- When agents interact, they can affect each others' evolution
- Co-evolutionary algorithms:
  - "models of social learning in which agents imitate strategies of other more successful agents."
- Genetic algorithm to learn $f(i,t), c(i,t), n(i,t)$
  - Initial weights
    - $f_{(i,0)} \sim |N(0, \sigma_f)|$
    - $c_{(i,0)} \sim |N(0, \sigma_c)|$
    - $n_{(i,0)} \sim |N(0, \sigma_n)|$
  - Population reproduces after every 5000 steps
  - If selected for reproduction, f, c, n or L_i is inherited.
  - Random mutations consist of f, c, n or L_i being changed

FTSE 100

Lag	Volume	Volatility	Order Signs	Returns
q=4	92.2	90.0	73.2	3.5
q=6	91.3	88.0	70.4	3.5
q=8	90.1	85.5	66.8	3.4
q=10	88.5	83.0	63.0	3.7

Results



Weight (f,c,n) Mean Values and Chartist Standard Deviation (σc)
Mean f
Mean c
Mean n
2
1.5
1
0.5
0
−0.5
−1
−1.5
2
3
4
5
6
7
8
Time
x 10^5

Where is the long memory?

Learning Phase

Table 5

LY Model Learning Phase. Percentages of runs with long memory for volume, volatility, order signs and returns at various time lags ranging from 4×50 (200 units of time) to 10×50 (500 units of time).

Lag	**Volume**	**Volatility**	**Order Signs**	**Returns**
q=4	70	74	13	2
q=6	69	71	11	2
q=8	68	69	11	3
q=10	68	68	10	3

Commitment Phase

Table 6

LY Model Commitment. Percentages of runs with long memory for volume, volatility, order signs and returns at various time lags ranging from 4×50 (200 units of time) to 10×50 (500 units of time).

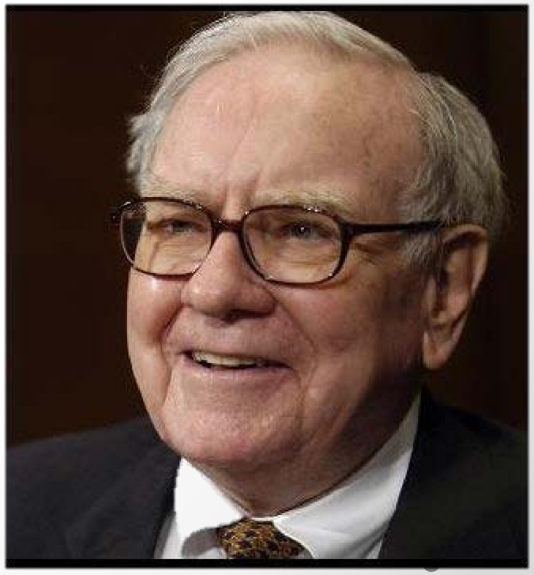
Lag	**Volume**	**Volatility**	**Order Signs**	**Returns**
q=4	2	3	2	4
q=6	2	2	2	4
q=8	2	2	3	4
q=10	3	1	3	4

Long memory

Lag	Volume	Volatility	Order Signs	Returns
q=4	2	3	2	4
q=6	2	2	2	4
q=8	2	2	3	4
q=10	3	1	3	4

Contrarian behavior

- Relaxed Constraint: Allow weights to be negative
- Two contrarian strategies
  - Negate the learnt trend
    - Example, contrarian chartist: $\hat{r}_{c_c(i,t,t+\tau)} = -\hat{r}_{c\ (i,t,t+\tau)}$
  - Do not follow current trend. Simply keep price constant
    - $\hat{r}_{f_c(i,t,t+\tau)} = \hat{r}_{n_c(i,t,t+\tau)} = 0$
- Contrarian variant
  - $\{\hat{r}_{c\ (i,t,t+\tau)}, \hat{r}_{f\ (i,t,t+\tau)}, \hat{r}_{n\ (i,t,t+\tau)}, \hat{r}_{f_c(i,t,t+\tau)}\ \hat{r}_{n_c(i,t,t+\tau)} \hat{r}_{c_c(i,t,t+\tau)}\}$



- Rayner et al

Results

Learning

Lag	Volume	Volatility	Order Signs	Returns
q=4	90	84	41	0
q=6	89	82	41	0
q=8	88	81	42	0
q=10	87	80	42	0

No Learning

Lag	Volume	Volatility	Order Signs	Returns
q=4	89	83	40	0
q=6	88	82	41	0
q=8	87	81	41	0
q=10	86	80	42	0

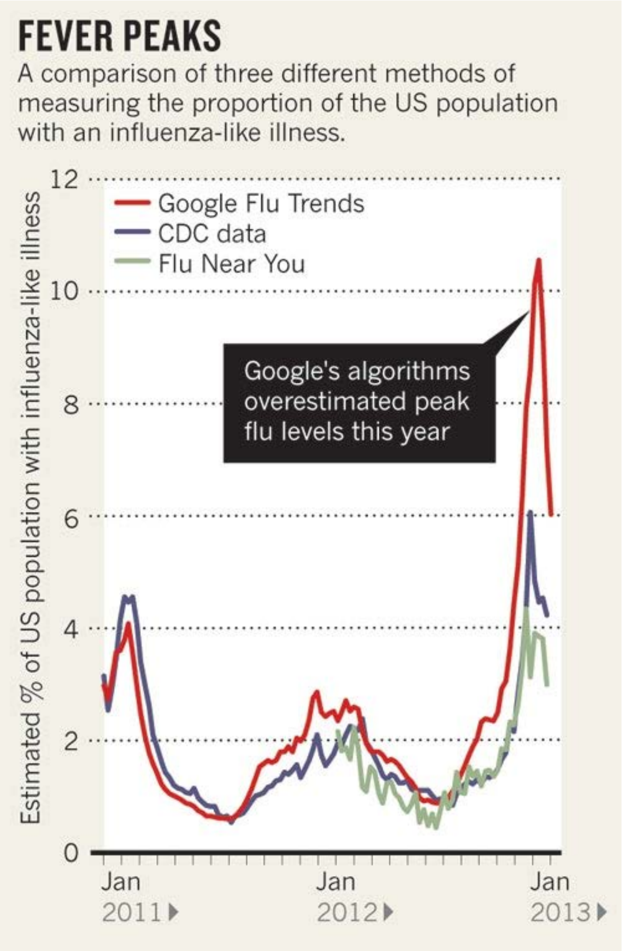
Potential Issues

- Feedback difficult to control
- Learned rules will be learned from static data
- Not suited for static, well-defined systems:
  - Image recognition
  - Voice recognition
- Some models may not be underpinned by rigorous mathematical proofs
  - Theory v practice
    - Rock/Paper/Scissors
  - Heuristics
- What level of recursion?
  - Do we really reach equilibriums
- Payoffs difficult to estimate in some domains
  - Non-monetary

Potential Issues, cont'd

- Features and design are very domain-specific
  - Not necessarily general-purpose

Google Flu Trends Prediction

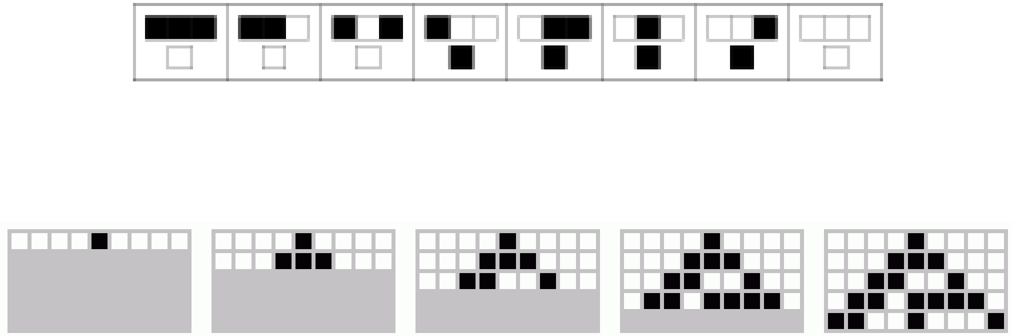


Additional Advantages

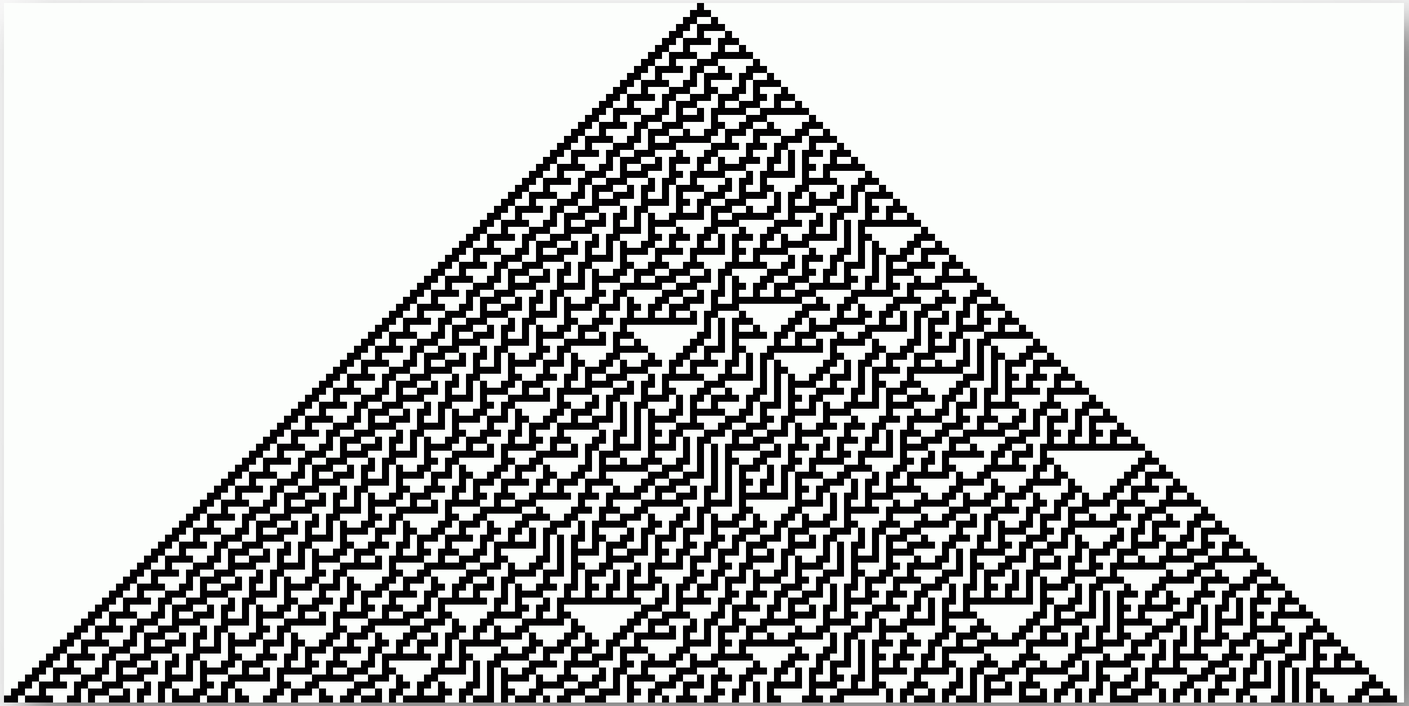
- Easier programming
  - Object-oriented
- Simpler rules/learning
- Closer to native architecture
- Adapts human knowledge/empirical results
  - "[T]he most useful learners are those that facilitate incorporating knowledge." Pedro Domingos

Simple rules over time

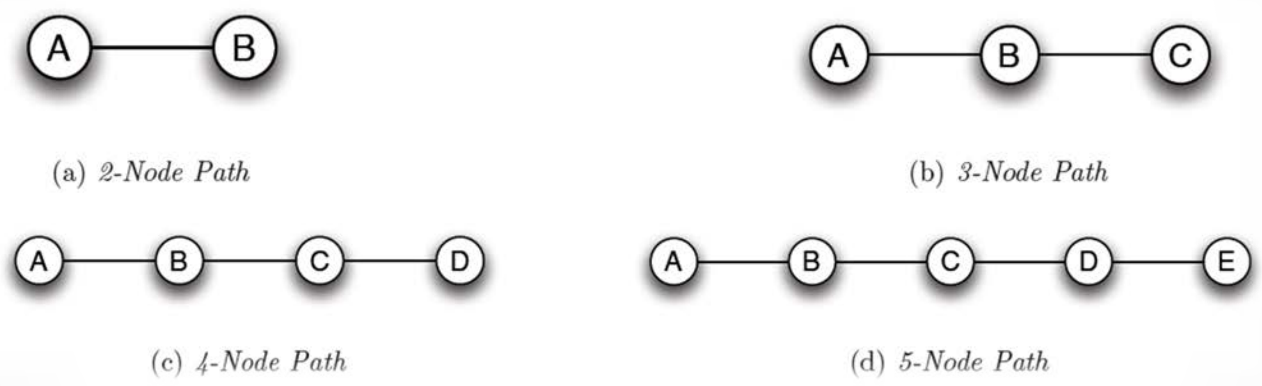
- Take the following simple rules;



Result



Bargaining: Reality vs Game Theory



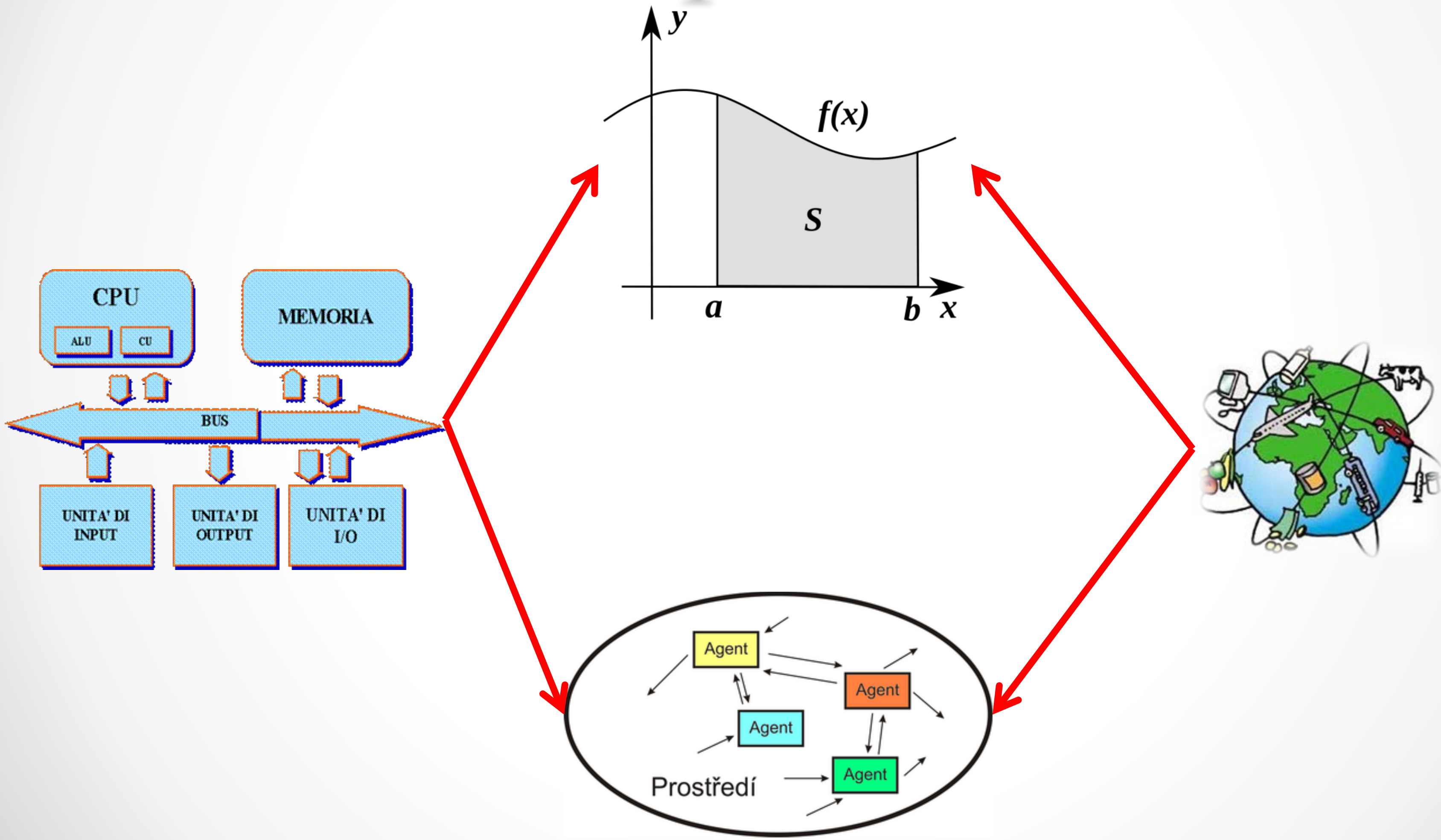
(a) *2-Node Path*

(b) *3-Node Path*

(c) *4-Node Path*

(d) *5-Node Path*

Which Representation?



Dichotomy

MAL Possibilities

- "…and they did work all manner of work of exceedingly curious workmanship."
  - Ether 10:27
- "…and became exceedingly rich in… machinery,…making all manner of tools of every kind…"
  - Jarom 1:8

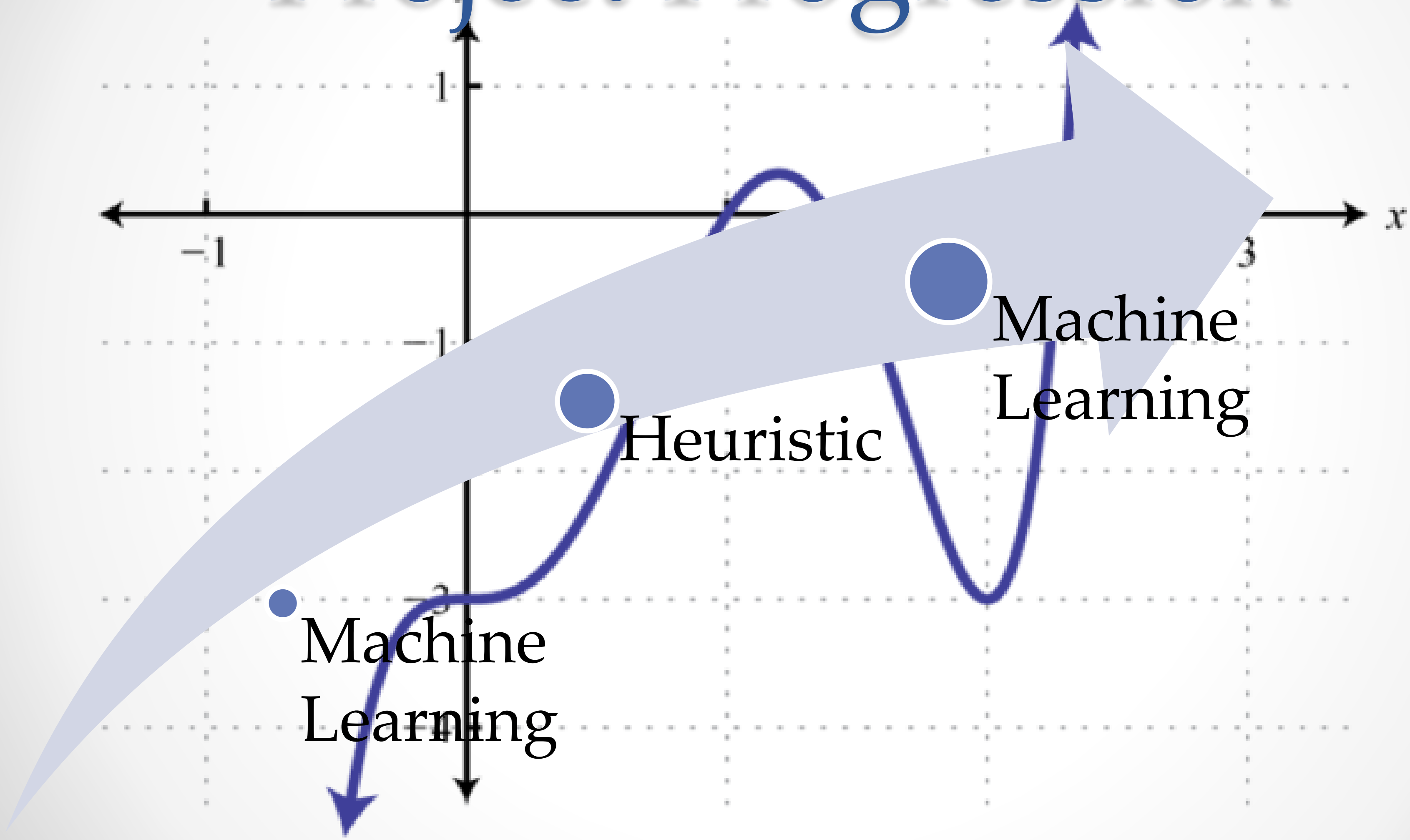
MAL Limits

- "Man can devise the most complex machines but cannot give them life or bestow upon them the powers of reason and judgment. These are divine gifts, bestowed only by God."
  - "Guided Safely Home" by President Thomas S. Monson

Upshot

- "There is no sharp frontier between designing learners and learning classifiers: rather, any given piece of knowledge could be encoded in the learner or learned from data. So machine learning projects often wind up having a significant component of learner design, and practitioners need to have some expertise in it."
    - Pedro Domingos

Project Progression



References

- Boyer, Carl B., and Uta C. Merzbach. A history of mathematics. John Wiley & Sons, 2011.
- Chaimontree, Santhana, Katie Atkinson, and Frans Coenen. "A multi-agent based approach to clustering: Harnessing the power of agents." Agents and Data Mining Interaction. Springer Berlin Heidelberg, 2012. 16-29.
- Czarnowski, Ireneusz. "Cluster-based instance selection for machine classification." Knowledge and Information Systems 30.1 (2012): 113-133.
- Domingos, Pedro. "A few useful things to know about machine learning." *Communications of the ACM* 55.10 (2012): 78-87.
- Easley, David, and Jon Kleinberg. Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press, 2010.
- Katz, Victor J. "A history of mathematics: an introduction." New York (1993).
- Panait, Liviu, and Sean Luke. "Cooperative multi-agent learning: The state of the art." *Autonomous Agents and Multi-Agent Systems* 11.3 (2005): 387-434.
- Rayner, Neil, Steve Phelps, and Nick Constantinou. "Learning is neither sufficient nor necessary: An agent-based model of long memory in financial markets." AI Communications (forthcoming) (2012).

References, cont’d

- Shoham, Yoav, and Kevin Leyton-Brown. Multiagent systems: Algorithmic, game-theoretic, and logical foundations. Cambridge University Press, 2008.
- Shoham, Yoav, Rob Powers, and Trond Grenager. "If multi-agent learning is the answer, what is the question?." Artificial Intelligence 171.7 (2007): 365-377.
- Siegfried, Tom. A Beautiful Math:: John Nash, Game Theory, and the Modern Quest for a Code of Nature. National Academies Press, 2006.
- Stone, Peter, and Manuela Veloso. "Multiagent systems: A survey from a machine learning perspective." Autonomous Robots 8.3 (2000): 345-383.
- Warnick, Sean. Presentation given in CS 513 Fall Semester, 2015.
- Wolfram, Stephen. A new kind of science. Vol. 5. Champaign: Wolfram media, 2002.
- Yang, Qiang, and Xindong Wu. "10 challenging problems in data mining research." International Journal of Information Technology & Decision Making 5.04 (2006): 597-604.
- http://www-history.mcs.st-andrews.ac.uk/Extras/Hilbert_Problems.html
- brainyquote.com