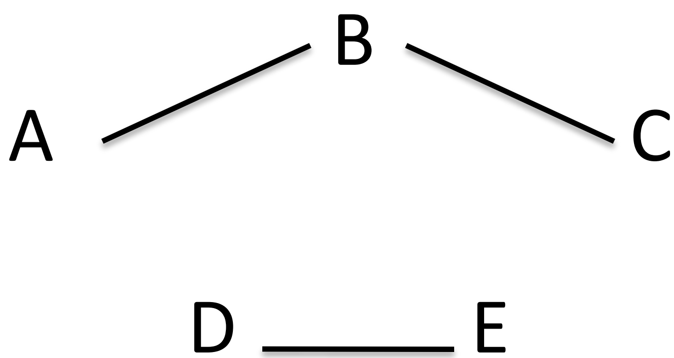
A Theory of Meaning:
Or How A Schema Reduces Complexity

Robert Axelrod

Complexity Conference on
“A Crude Look at the Whole”
Nanyang Technological University
Singapore
March 4-6, 2013

Example: A Balanced Schema

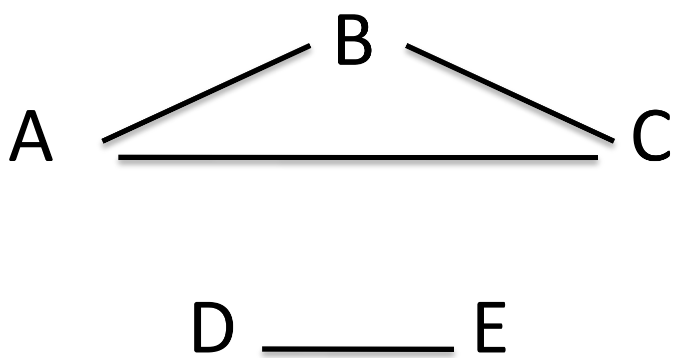
1. Friends of friends are friends.
2. Friends of enemies are enemies.
3. Enemies of enemies are friends.



You know A likes B, B likes C, and D likes E.

Example: A Balanced Schema

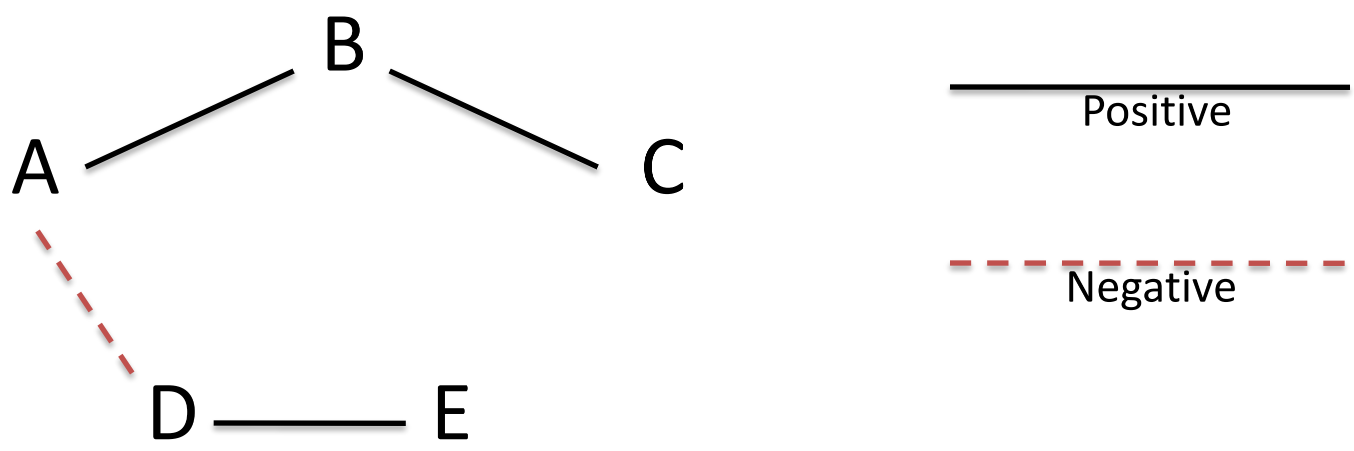
1. Friends of friends are friends.
2. Friends of enemies are enemies.
3. Enemies of enemies are friends.



You can infer that A likes C.

Example: A Balanced Schema

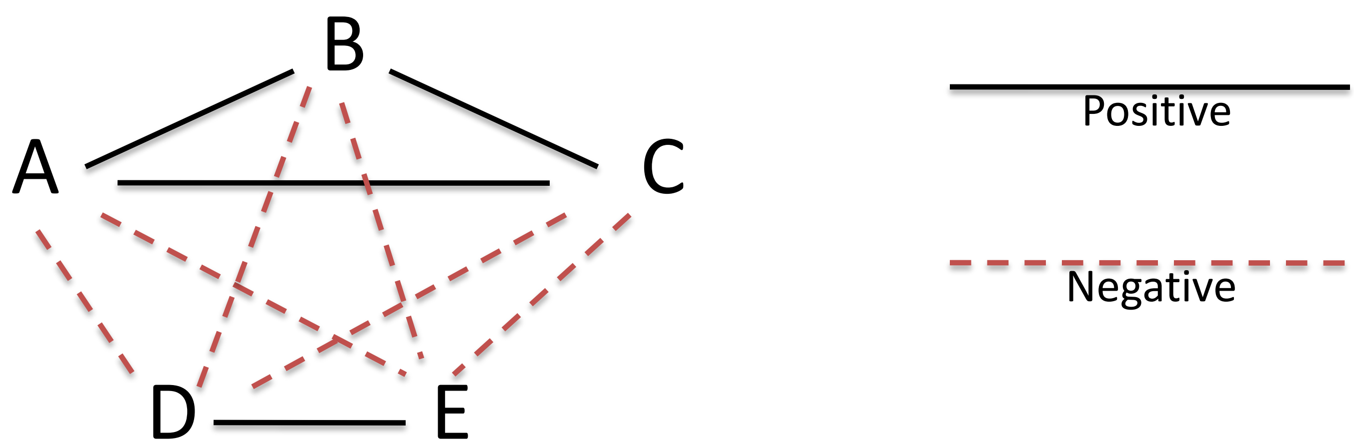
- You know: A likes B, B likes C, and D likes E
- You learn that A dislikes D.



Now you can infer the remaining 6 links.

Example: A Balanced Schema

- You know: A likes B, B likes C and D likes E
- You learn that A dislikes D.



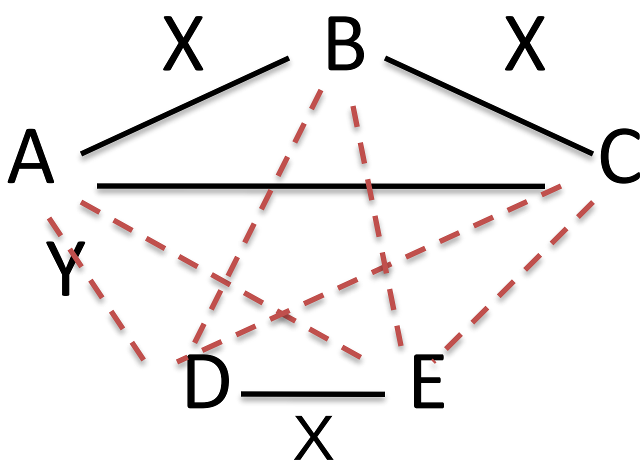
Now you can infer the remaining 6 links.

Meaning of news Y={AD-} given balance and prior info X = {AB+, BC+, DE+} is 6 bits

$$M_{bal}(Y|\ X) = Info_{bal}(X\ U\ Y) - Info_{bal}(X) - Info_{bal}(Y)$$

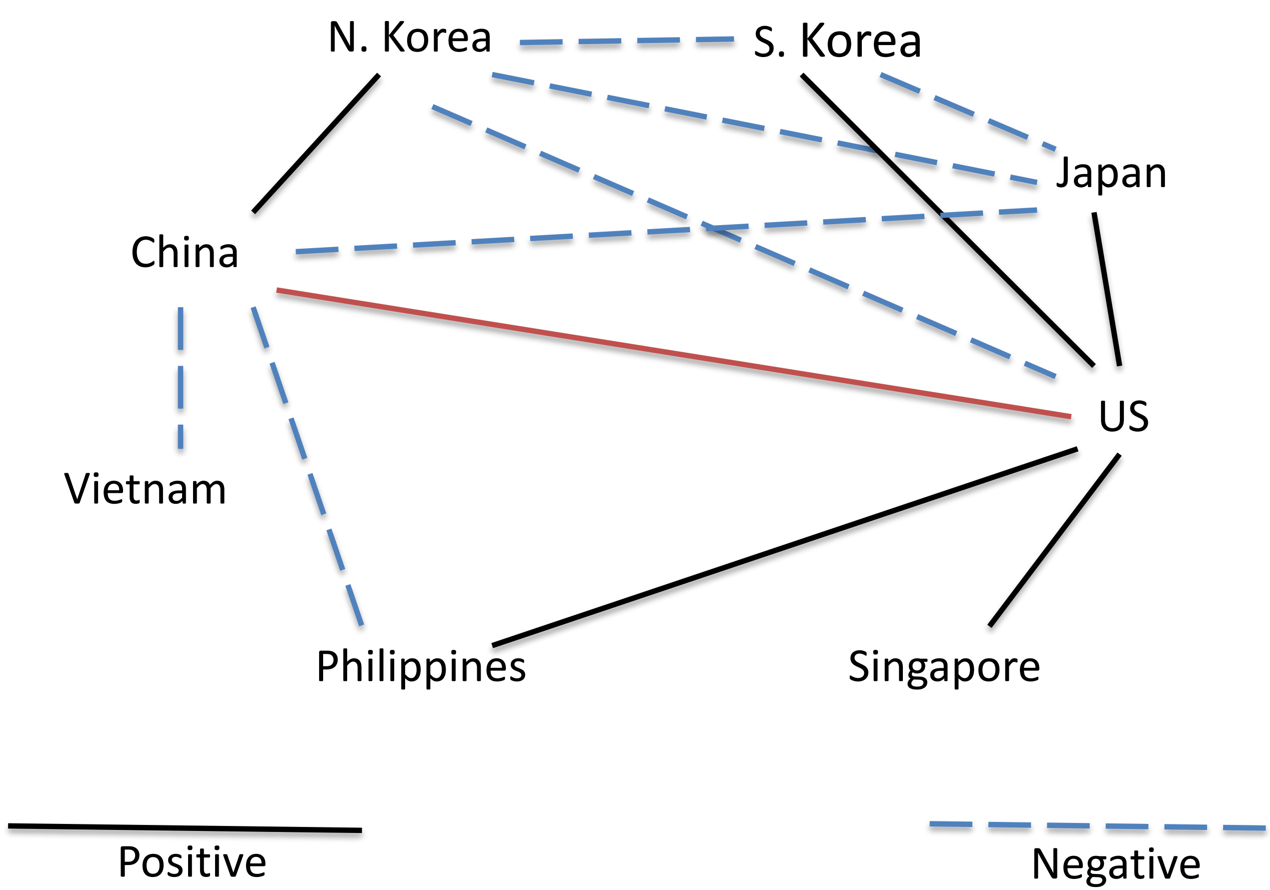
$$= 10 \text{ bits} - 3 \text{ bits} - 1 \text{ bit}$$

$$= 6 \text{ bits}$$



Related to Gell-Mann and Lloyd's Effective Complexity.

Relations Among Selected Pacific Rim Countries



Some Definitions

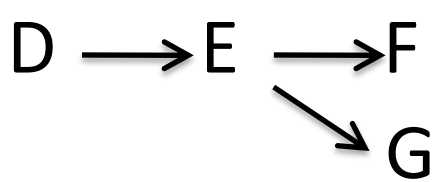
- State space: allowable configurations
  - Example: any graph (network)
    - with a set of nodes, and
    - a positive or negative link between each pair of nodes
- Schema: a subset of the state space
  - Example: *balanced* graphs

Another Schema: Rank Order

1. A > B or B > A
2. If A > B, and B > C then A > C.

Example 1: When A > B > C is known, meaning of
C > D is A > D and B > D. A > B > C > D

Example 2: When D > E > F is known, meaning of
E > G is D > G. But don't know if F > G.



Uses of a Schema

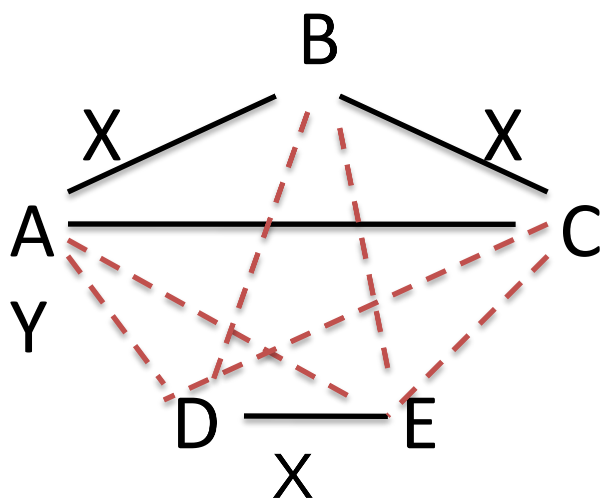
1. Infer Missing Data
2. Correct Errors
3. Predict Changes
4. Economize on Memory
5. Evaluate Source Credibility
6. Measure Stress and Goodness of Fit
7. Focus Attention to Resolve Ambiguity

1. Infer Missing Data

Given X = A likes B, and B likes C, and

Y = A dislikes D

The balance schema allows one to infer the other six relationships.



2. Error Correction

Two balanced graphs differ by at least n-1 links.

- Example: ABC versus DEFGH where n = 8.
- If C moves from ABC to DEFGH, C will change all 7 of its links.
- If more than one nodes changes sides, more than 7 links change.

So up to (n-1)/2 errors can be corrected.

- Example: An unbalanced graph of 8 nodes that is 1, 2, or 3 links away from a balanced graph is at least 4 links away from any other balanced graph.
- Thus 1, 2 or 3 errors can be corrected.
- For n=20, up to 9 errors can be corrected since 9 < (20-1)/2.

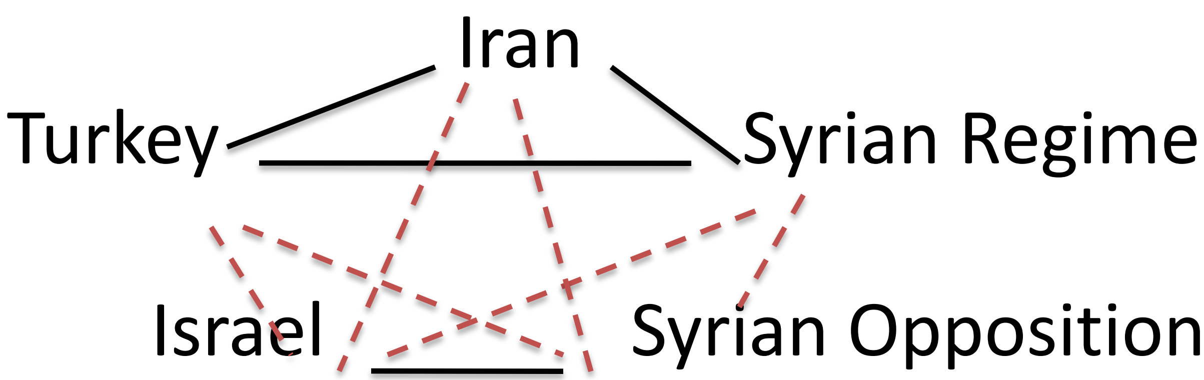
3. Predict Changes

- Empirical claim: An unbalanced graph is attracted to one of the closest balanced graphs. (If n is even, this is unique.)
  - Example:
    - AB vs. DEFGH, but
    - C likes A but not B, and likes D but not E, F, G and H.
    - Prediction: C will come like A, and dislike D, making the balanced graph ABC vs. DEFGH.
- So changes in links can be predicted.

3. Predict Changes, cont.

2011: Turkey sides with Iran and Syrian Regime.

Israel is against Syrian Regime, so sides with Syrian Opposition.



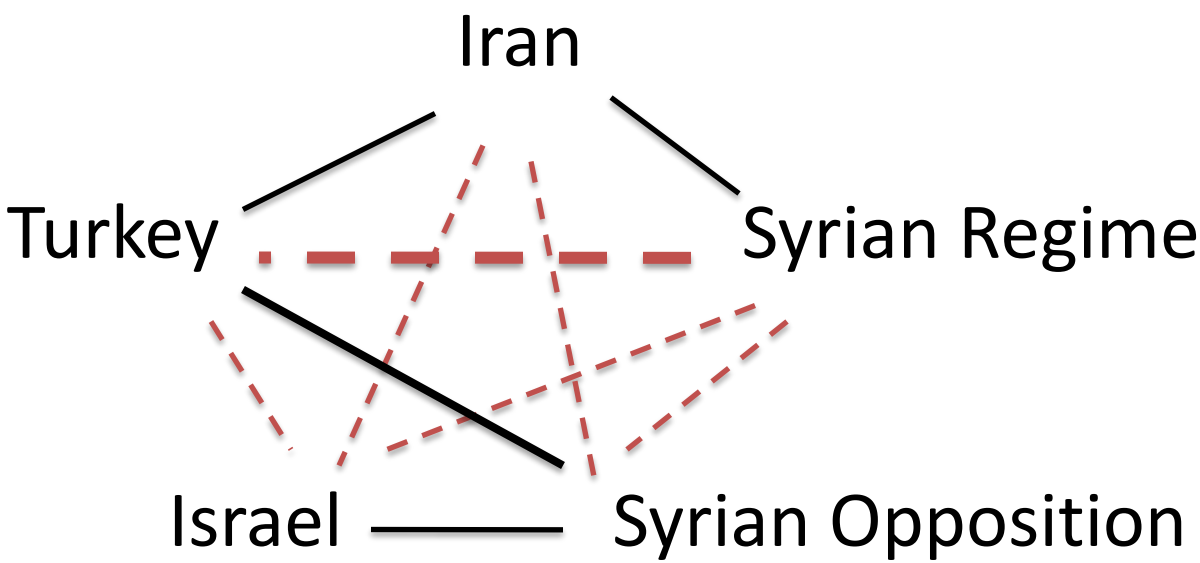
.

.

.

3. Predict Changes, cont.

2012. Turkey opposes Syrian Regime and favors Syrian Opposition.

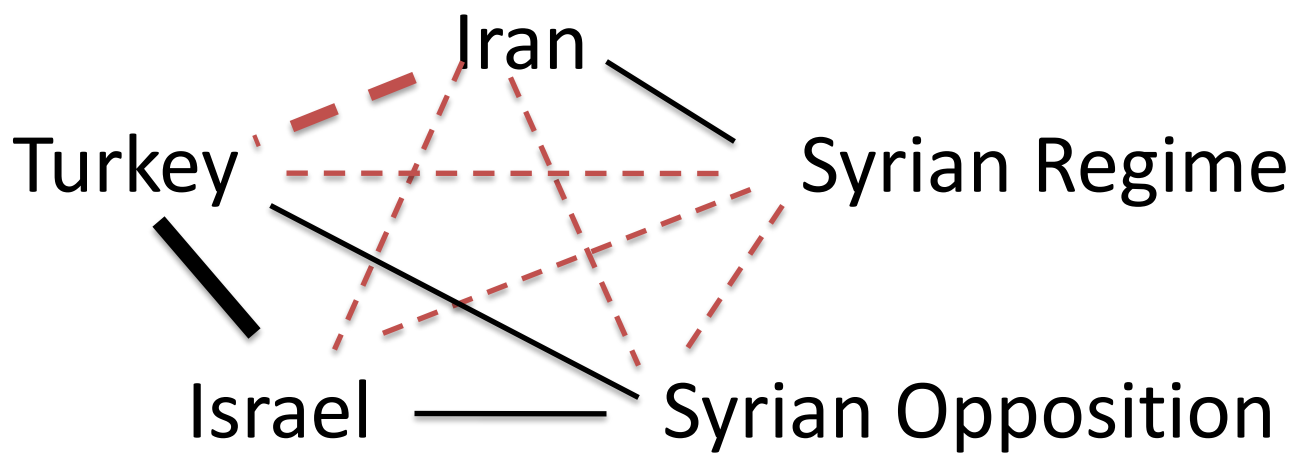


.

Prediction: Turkey will dislike Iran and like Israel.

Predicting Change, cont.

Feb. 24,2013. Reuters. “Despite the widespread belief that ties between Israel and Turkey are virtually non-existent, in fact, Israeli and Turkish officials have held a series of meetings, the most recent three weeks ago, according to Israeli officials who confirmed media reports in both countries.”



These two changes would achieve a new balance.

4. Economize on Memory

- A graph of n nodes has n(n-1)/2 links.
  - Example: 8 nodes have 8*7/2 = 28 links.
- A balanced graph of n nodes requires only n-1 links to infer the rest (provided the n-1 links connect all the nodes).
  - Example: 8 nodes need 7 links, e.g. all of A's likes and dislikes. This compared to 28 links in general.
  - Example: 20 nodes need 19 links if balanced, but 190 if not. This saves 90% on memory.

5. Evaluate Source Credibility

- Use the error correction method to identify links that need changing to get to the nearest balanced graph.
- Then reduce the credibility of the source(s) that provided the “erroneous” information.
- This can improve accuracy in the future.

6. Measure Stress

- Def. Stress in a graph is the distance to the nearest schema (e.g. the nearest balanced graph).
- Stress can be used to evaluate goodness of fit of a graph to a schema.
  - Example: Let G be the graph that has three groups of friends all of whom dislike those in other two groups. Then G is not balanced, but it does fit a 3-group schema.
- Prediction (reformulated from before): For a given schema, links will change to minimize stress.
  - Axelrod's Maxim: "Politics <u>minimizes</u> the strangeness of bedfellows."

7. Focus Attention to Resolve Ambiguity

- If a graph is nearly equidistant to two balanced graphs, one can identify the links that would need to change to go toward one or the other.
- These links can be studied to resolve the ambiguity about which way the graph is likely to go (or which observations are erroneous).

Beliefs and Reality Can Interact To Sustain a Schema

Three reasons to expect the world fits a balanced schema:

1. Changes in the world move toward balance (stress reduction).
2. A balanced schema is cognitively efficient.
3. Belief in a schema can be self-fulfilling, because reasons 1 and 2 can reinforce each other.

Various Schemas

- M clusters (e.g., balance is 2-cluster)
- Rank Order
- Multidimensional (e.g., 2 dim mileage chart)
- Trees (e.g. organizational chart; tree of life)
  - Default hierarchies: e.g., dogs have properties of mammals
- Etc.

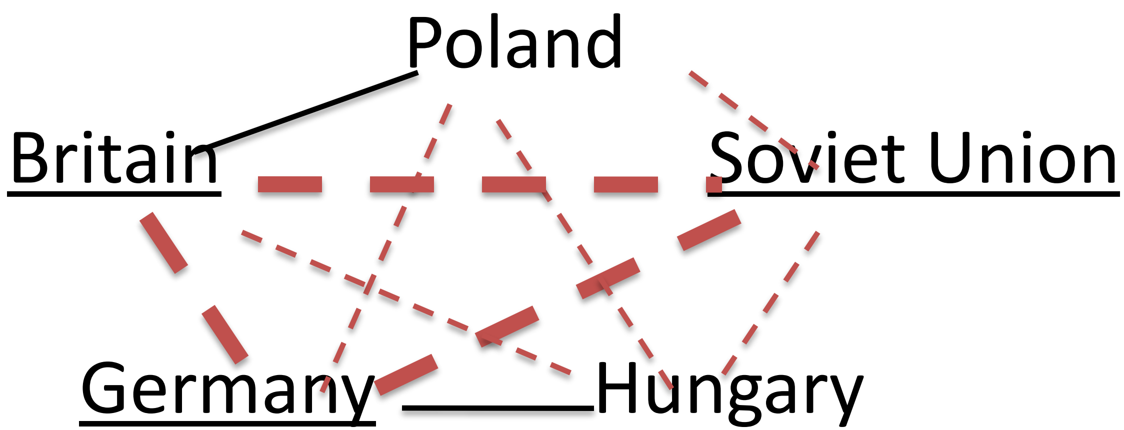
Schemas differ on their robustness (the size of basins of attraction).

Choosing Between Schemas

- Suppose cost = # incorrect beliefs
- Suppose there are priors over schema.
  - Example: 80% balance, 20% 3-cluster
- Then for a given (complete) observed graph, one could choose the schema and (nearest) graph that minimizes expected cost.
- This gives “optimal beliefs.”

Alignment in World War II in Europe

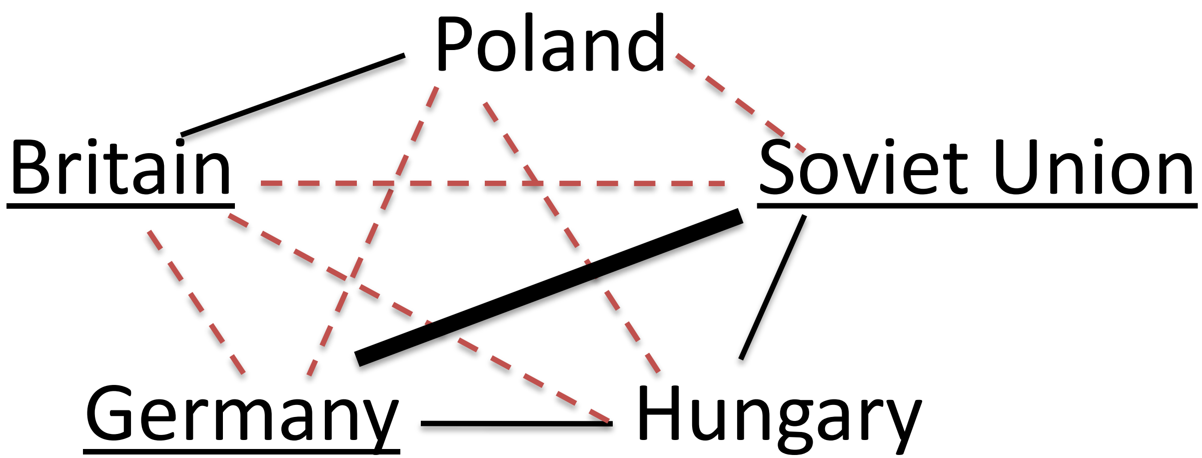
- N =17 countries involved
- Before the war, situation is unbalanced



Example: Britain is anti-Nazi and anti-Communist, but Germany and SU hate each other too.

Alignment in World War II in Europe

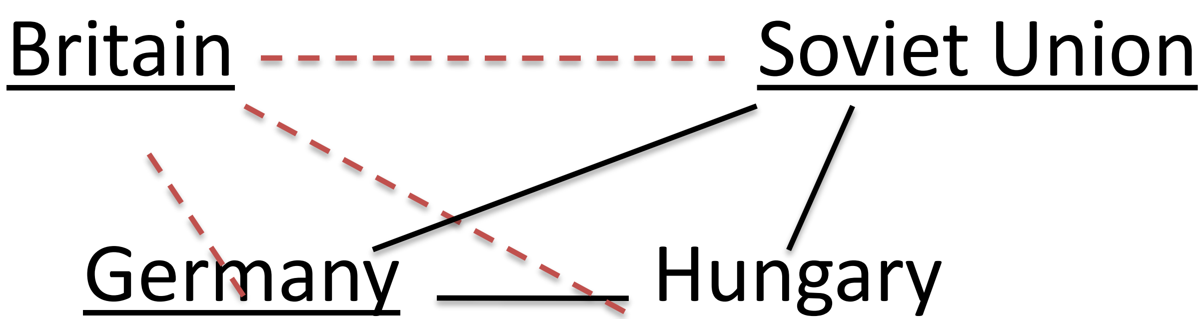
- In 1939 Hitler and Stalin sign a pact.
- Germany (and Hungary) align with SU.



.This is now balanced.

Alignment in World War II in Europe

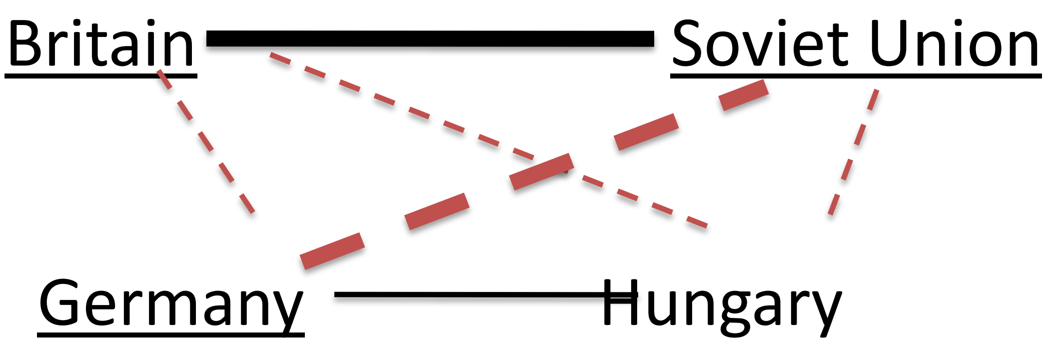
- In 1939 Hitler and Stalin sign a pact.
- Germany (and Hungary) align with SU.



.This is now balanced. But Germany and Soviet Union still hate each other.

Alignment in World War II in Europe

- 1941. Germany invades Soviet Union
- Churchill explains the meaning of this news. "If Hitler invaded hell I would make at least a favourable reference to the devil in the House of Commons."



This too is balanced. And it has less stress.

Predicting Alignment

- N =17 countries involved. There are n(n-1)/2 = 136 links.
  There are $2^{(17-1)}$ = 65,536 balanced configurations.

- But now links not binary. A link is the propensity of two countries to get along with each other. Propensities are estimated by:
  - Similarity of religion, Similarity of government type, Previous war, Ethnic conflict, and Border issue

- A country's stress in a given configuration =
  (weighted propensities with countries on its own side)
  – (weighted propensities with countries on the other side).
  Weight of a country is its national capability (demographic, industrial, military)

- This gives a configuration space: points in this space are balanced graphs.

- Premise: A country will change sides if it would reduce its stress. This can be regarded as movement in the configuration space.

- Start with a random alignment, and let countries change sides until no one wants to change any more. Prediction: local minima.

Prediction: One of Just Two Configurations

TABLE 1 *The Two Configurations Predicted for the Second World War in Europe**

		Configuration 1 Alignment 1	Alignment 2
Configuration 2	Alignment 1	Britain (7.45) France (5.32) Czechoslovakia (1.15) Denmark (0.20)	Germany (11.49) Italy (4.03) Poland (1.83) Romania (0.78) Hungary (0.45) Portugal (0.27) Finland (0.19) Latvia (0.13) Lithuania (0.10) Estonia (0.06)
	Alignment 2	Soviet Union (15.01) Yugoslavia (0.59) Greece (0.35)	(None)
Nearest empirical match†		Allies (and those invaded by Germany)	Axis (and those invaded by the Soviet Union

$p < 0.005$

*The size is shown in parentheses, in terms of percentage of world capabilities. The predicitons are based upon 1936 data.

†In Configuration 1, only Poland and Portugal are wrong.

Energy (Stress) Landscape for State Space of 2^{16} Configurations

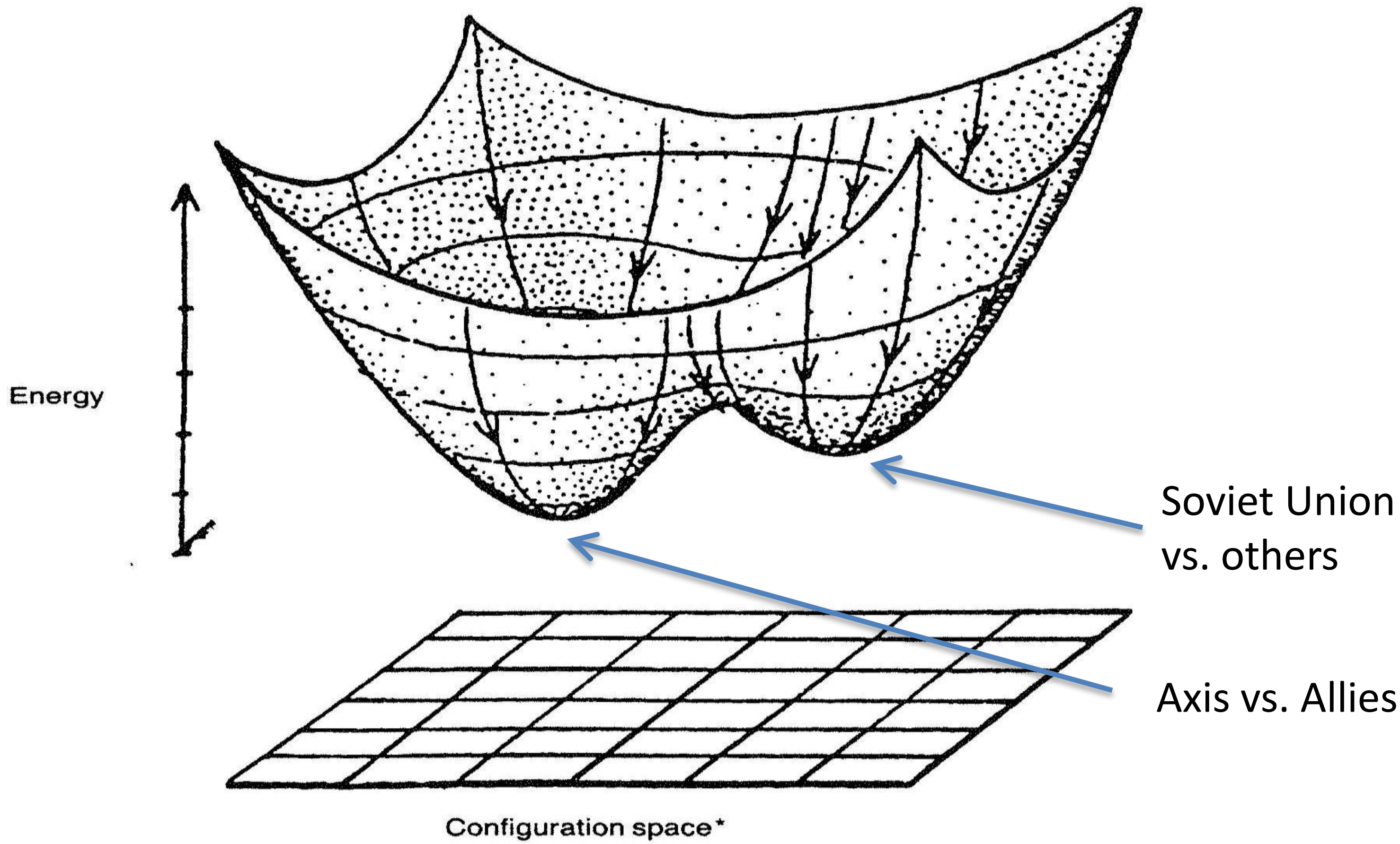


Fig. 1. A landscape with two local optima
Note: Adapted from Abraham and Shaw, *Dynamics*, Part 2; used with the permission of Ariel Press.
*The configuration space is an *n*-dimensional binary hypercube. The hypercube has one dimension for each firm indicating which of the two possible alliances that firm is in.

Setting Standards for Unix

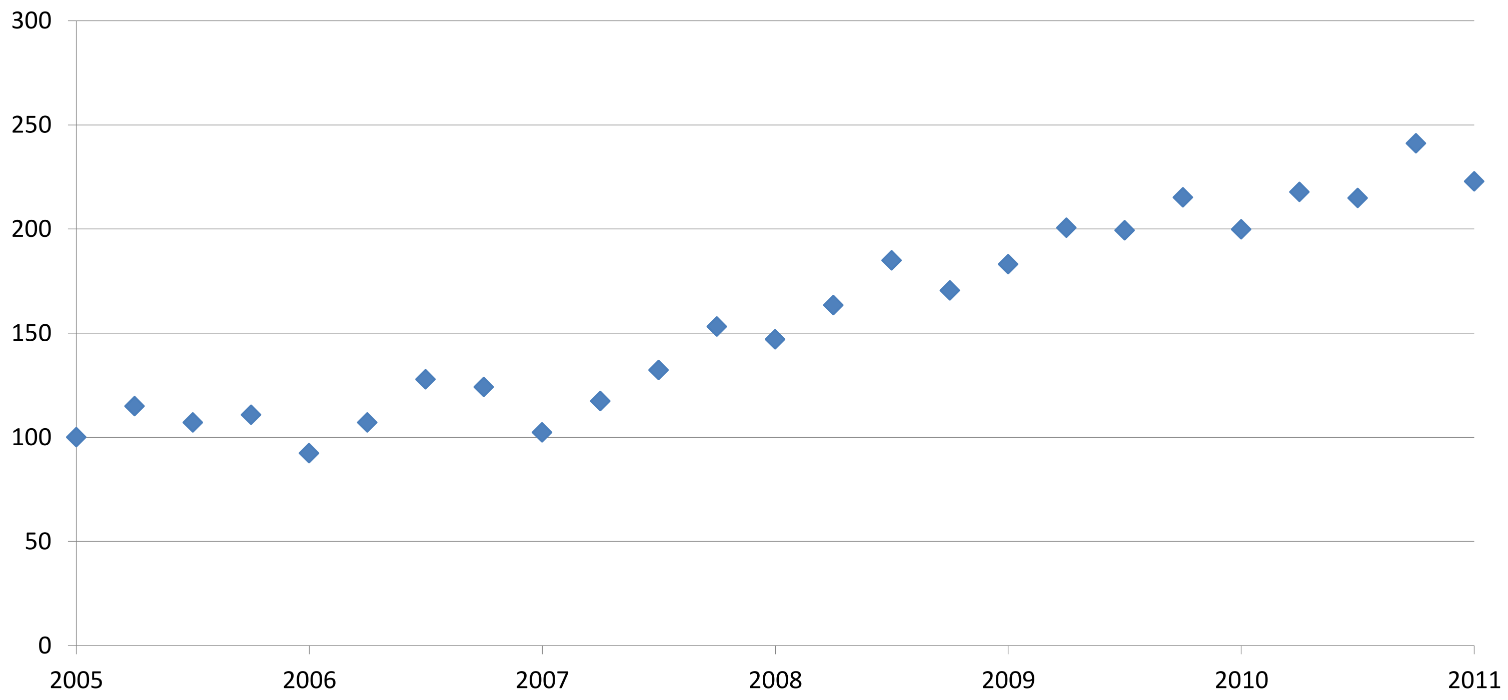
TABLE 2 *The Two Configurations Predicted for Unix Alliances**

	Configuration 1 Alignment 1	*Configuration 1* Alignment 2
Alliance 1	Sun (28.9, S) Prime (1.0, S)	DEC (20.0, G) HP (11.5, G)
Configuration 2		
Alliance 2	AT&T (28.5, G) IBM (3.8, G)	Apollo (21.2, S) Intergraph (4.4, S) SGI (4.4, S)
Nearest empirical match†	Unix International	Open Software Foundation

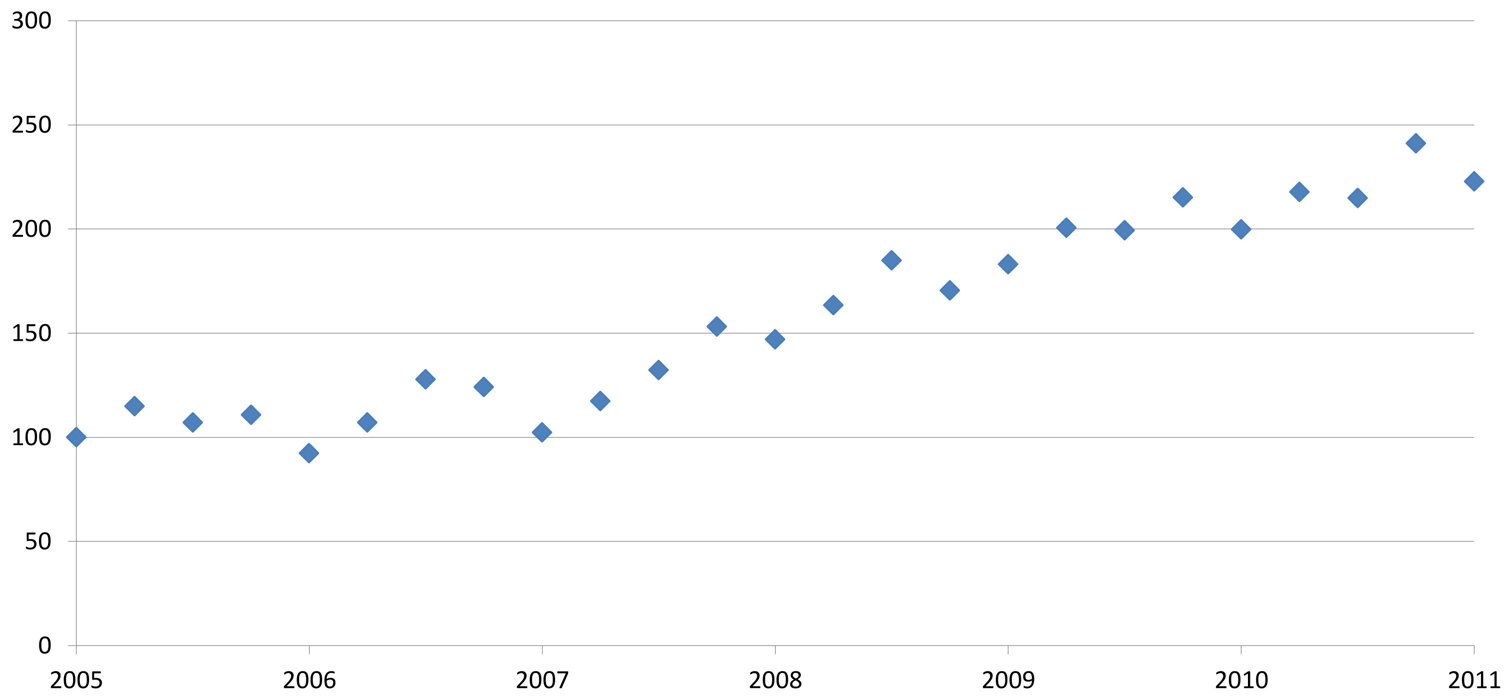
*Size is shown in parentheses, along with whether the firm was a computer generalist (G) or technical workstation specialist (S). All firms had a design orientation.

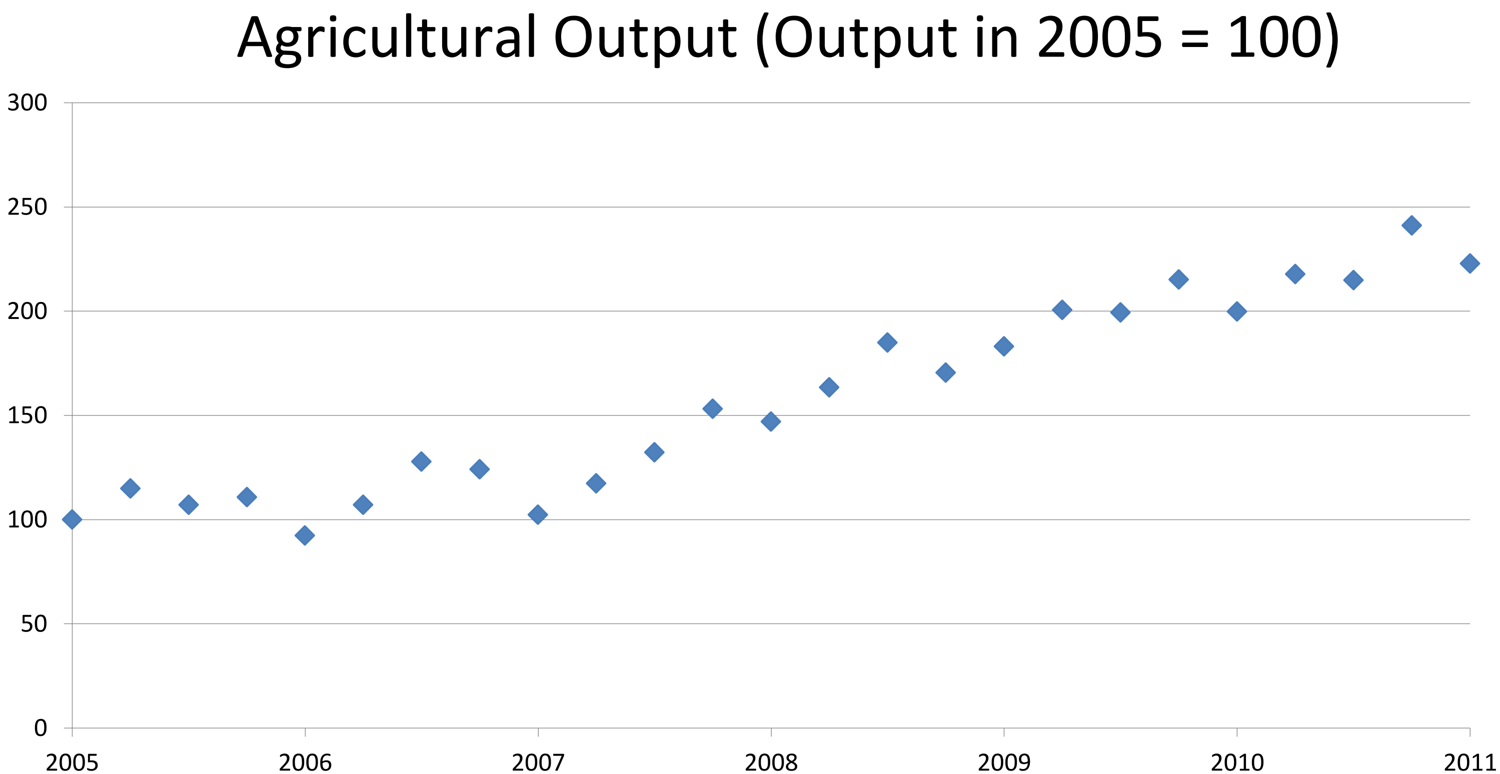
† In Configuration 1, only the IBM prediction is wrong.

Second Type of Application: Schemas For Curve Fitting



Housing Construction (2005 = 100)





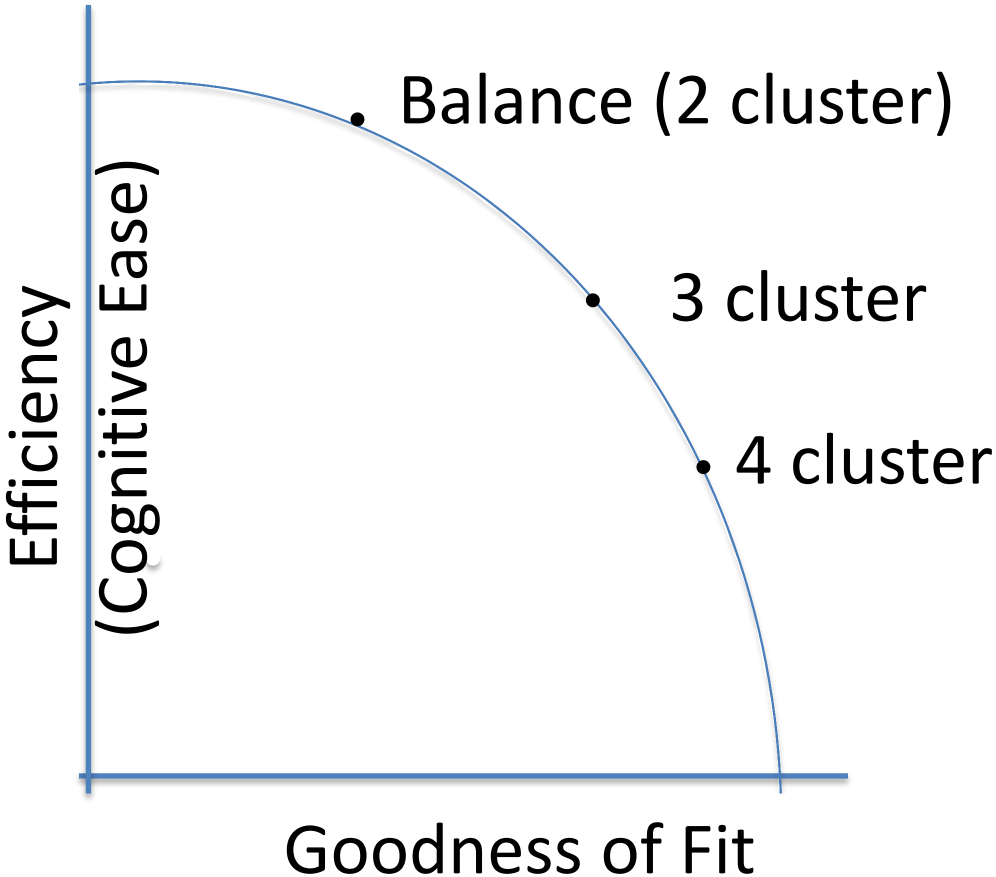
Exponential and seasonal with noise: $y(t) = y(t-1)*(1+ g/4) + s + w$

g = 3 %, quarterly growth rate

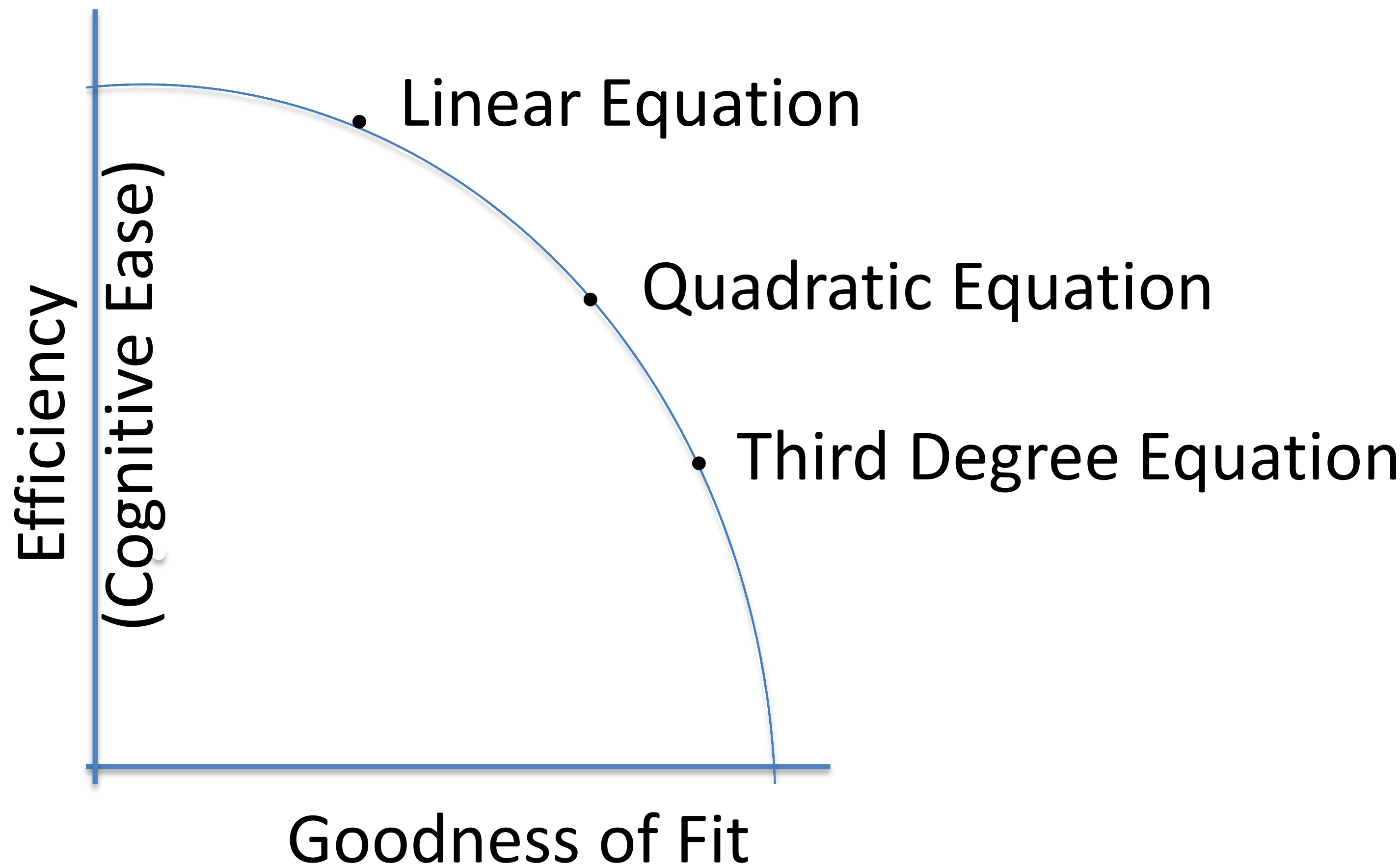
s = 12 seasonal component which gives a boost in Q2, and a decrement in Q4

w weather: random number, uniformly chosen between - 20 and 20

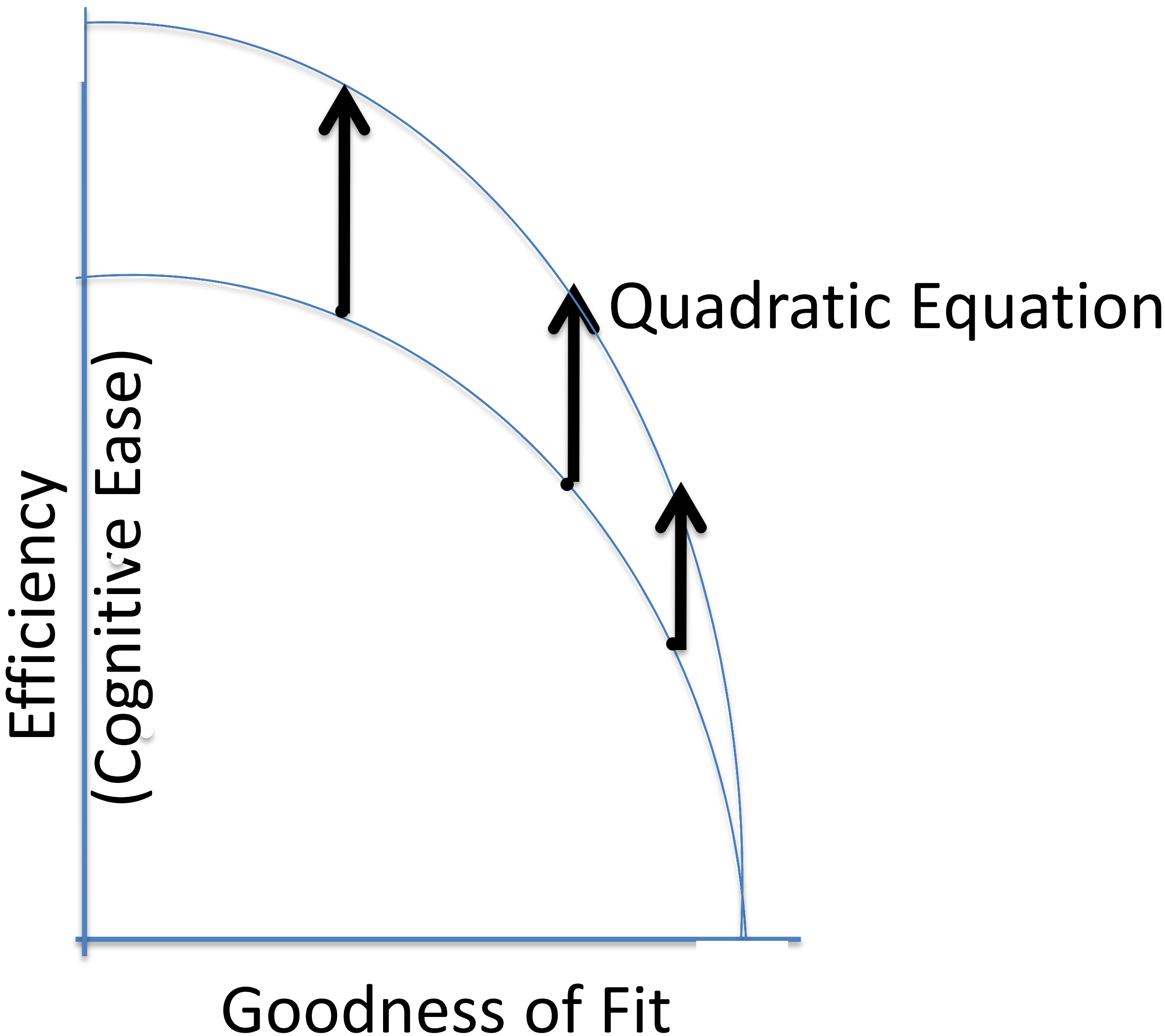
Choice of a Schema Trades Off Efficiency and Goodness of Fit



Choice of a Schema Trades Off Efficiency and Goodness of Fit



A Course in Algebra Helps



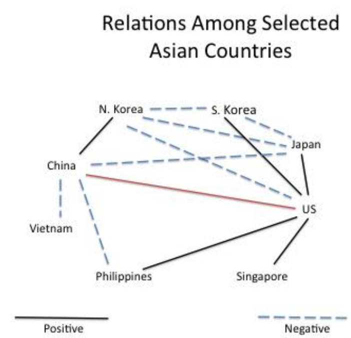
Review.
Uses of a Schema Reduces Complexity:

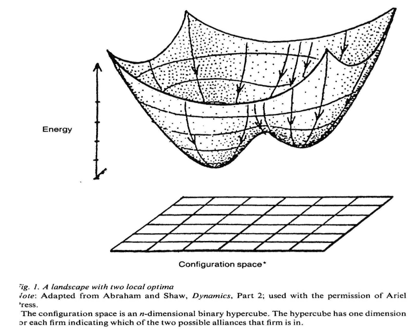
1. Infer Missing Data
2. Correct Errors
3. Predict Changes
4. Economize on Memory
5. Evaluate Source Credibility
6. Measure stress and goodness of fit
7. Focus attention to resolve ambiguity

Third Type of Application: Interpreting Pictures.
A Typical Problem in Artificial Intelligence
The Facial Schema with Two Interpretations

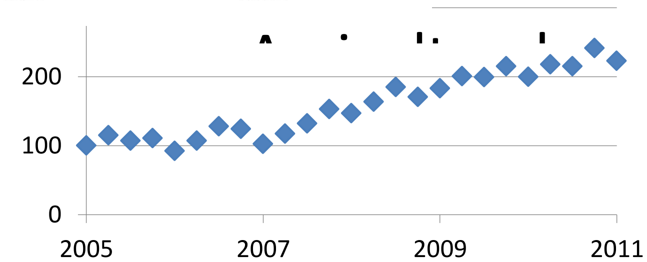
Conclusion:
Schemas Can Help Reduce Complexity

1. Network of relations
    - Balance, m-clusters
    - Rank Order





2. Curve Fitting



3. interpreting Pictures

Thank you.

How Many Legs Does This Have?

Is There Something Else Here?

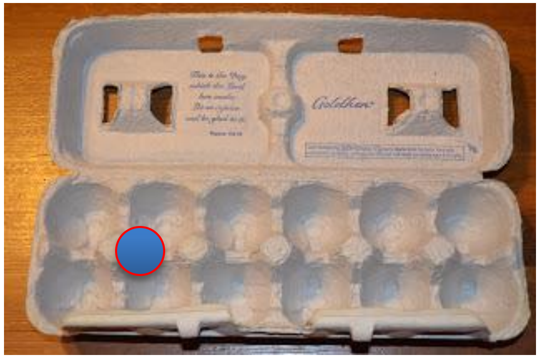
Questions and a Premise

- How can one make sense of a novel event?
  - How do people actually do it?
  - How should people do it?
  - How can artificial intelligence do it?
- Premise:
  - An event is experienced as information about particular kinds of relationships among particular actors.
    - Example: A likes B.

Time Scales of Graph Dynamics

Assume links change at random, with bias toward nearest balanced graph.

- 1. Very short time scale: at observation



Time Scales of Graph Dynamics

Assume links change at random, with bias toward nearest balanced graph.

- 1. Very short time scale: at observation
- 2. Short time scale: move toward nearest balanced graph (within the basin of attraction).
-

Time Scales of Graph Dynamics

Assume links change at random, with bias toward nearest balanced graph.

- 1. Very short time scale: at observation
- 2. Short time scale: move toward nearest balanced graph (within the basin of attraction).
- 3 Medium time scale: at or very near the nearest balanced graph, i.e.

  the local minimum.

Time Scales of Graph Dynamics

Assume links change at random, with bias toward nearest balanced graph.

- 1. Very short time scale: at observation
- 2. Short time scale: move toward nearest balanced graph (within the basin of attraction).
- 3 Medium time scale: at or very near the nearest balanced graph, i.e. the local minimum.
- 4. Long time scale: at or very near *any* balanced graph: a slow random walk among neighboring balanced graphs.

Time Scales of Graph Dynamics

Assume links change at random, with bias toward nearest balanced graph.

- 1. Very short time scale: at observation
- 2. Short time scale: move toward nearest balanced graph (within the basin of attraction).
- 3 Medium time scale: at or very near the nearest balanced graph, i.e. the local minimum.
- 4. Long time scale: at or very near *any* balanced graph: a slow random walk among neighboring balanced graphs.

Time Scales of Graph Dynamics

Assume links change at random, with bias toward nearest balanced graph.

- 1. Very short time scale: at observation
- 2. Short time scale: move toward nearest balanced graph (within the basin of attraction).
- 3 Medium time scale: at or very near the nearest balanced graph, i.e. the local minimum.
- 4. Long time scale: at or very near *any* balanced graph: a slow random walk

among neighboring

balanced graphs.