USC Viterbi School of Engineering

USC

Evaluating Posterior Probabilities of Mental Models

Jonathan Y. Ito

David V. Pynadath

Stacy C. Marsella

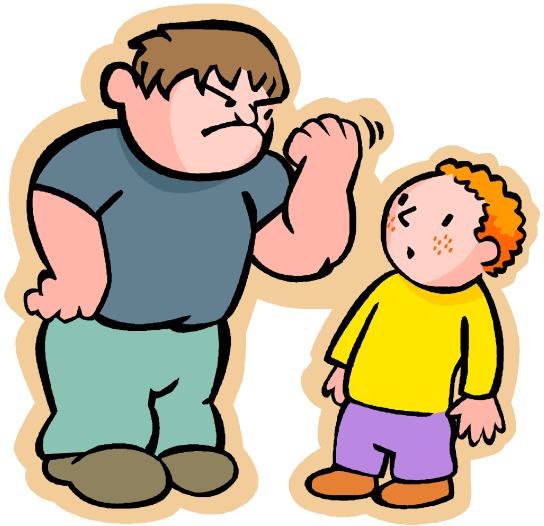
ISI Information Sciences Institute

Schoolyard Scenario

Onlooker

Teacher

Bully



Victim

USC

ISI
Information Sciences Institute

Bully's Thought Process

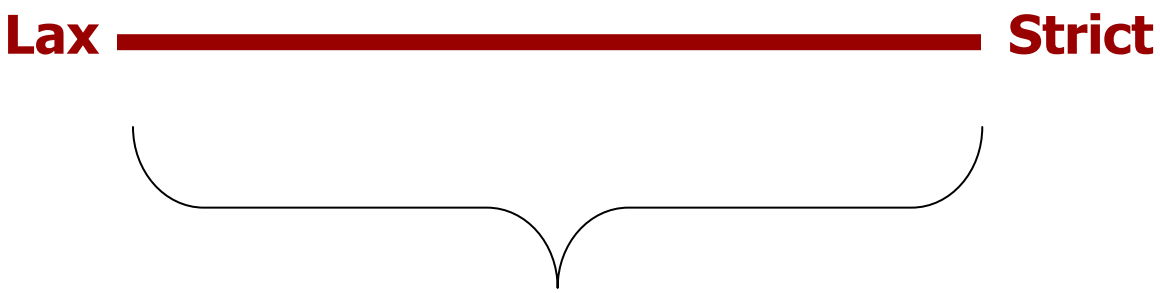
Bully observes teacher punishing class

Bully updates estimation of teacher to very strict

USC

ISI
Information Sciences Institute

Teacher's Mental Model Space

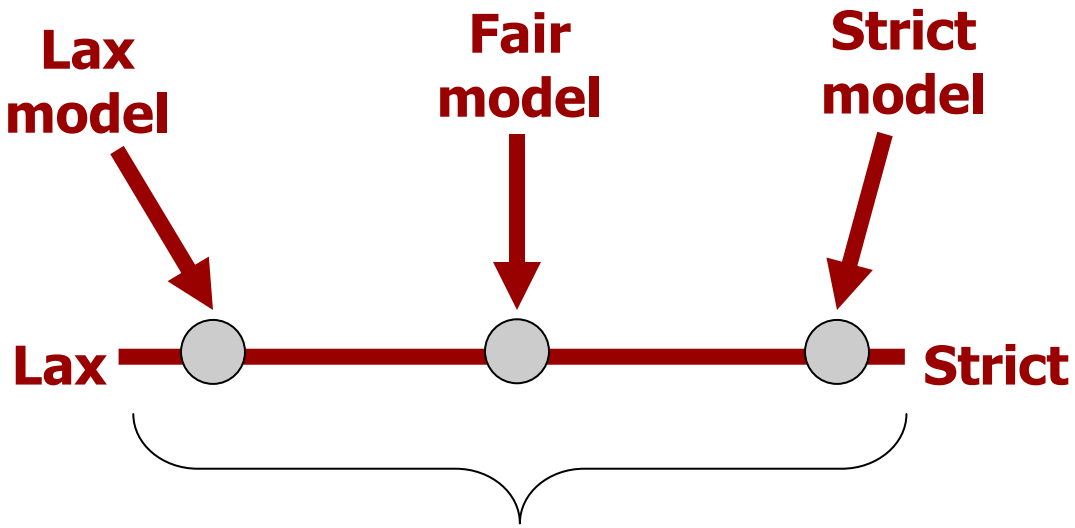


USC

ISI
Information Sciences Institute

What Does the Bully Consider?

- **Continuous space of mental models is too big!**
  - Must choose a discrete number of mental models to partition the space

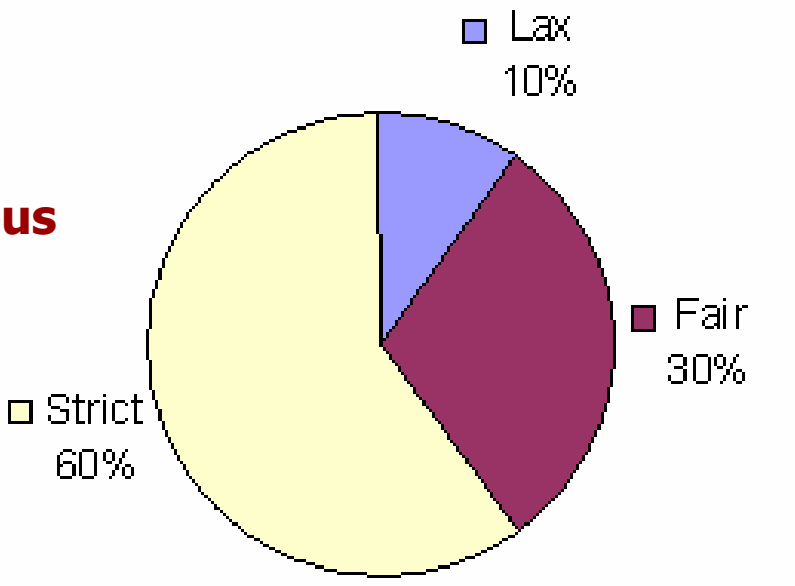


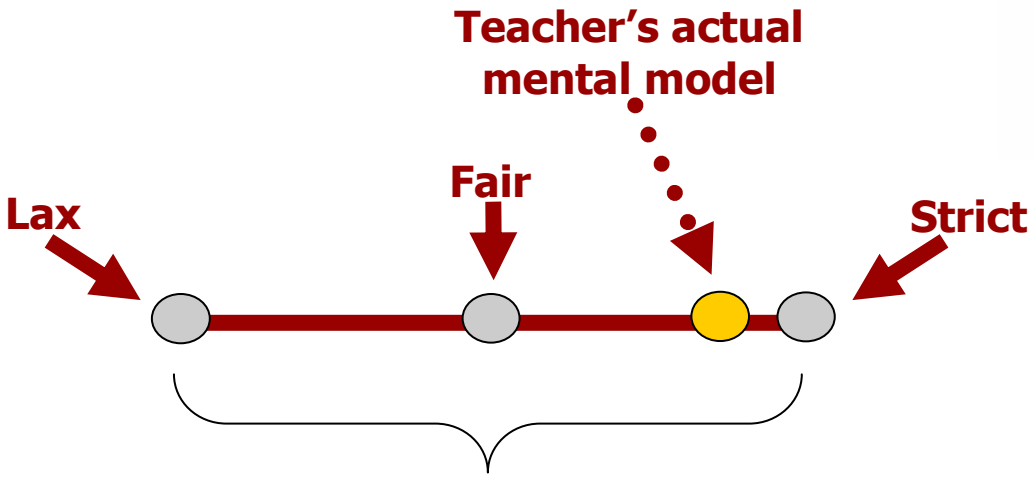
USC

ISI
Information Sciences Institute

What Does the Bully Believe?

- Choosing 1 mental model is too coarse
- Use a distribution instead!
- Can't have a distribution over continuous space



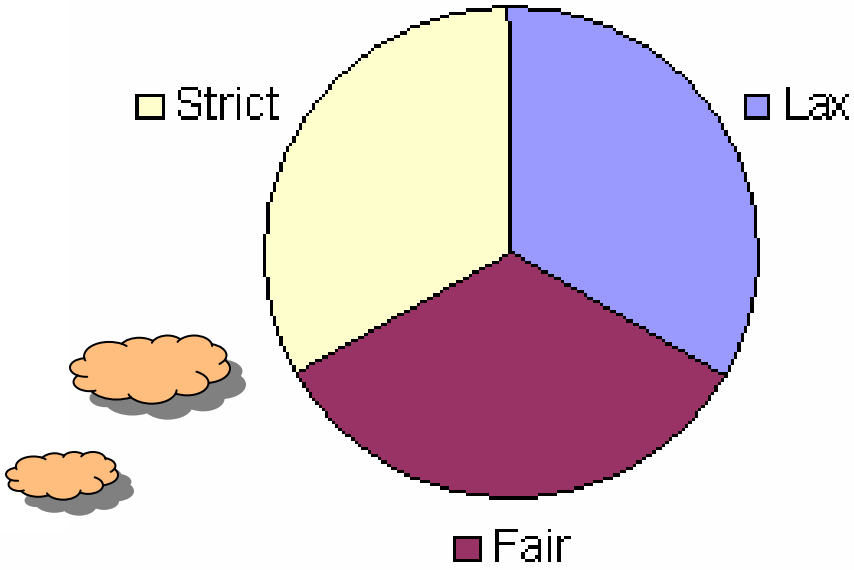


USC

Information Sciences Institute

Example – Initial Beliefs

- **Bully has some initial estimation of teacher's mental models**



USC

ISI
Information Sciences Institute

Example – Actions and Observations

- Bully takes and observes actions in the world

Bully picks on Victim

Onlooker laughs at Victim

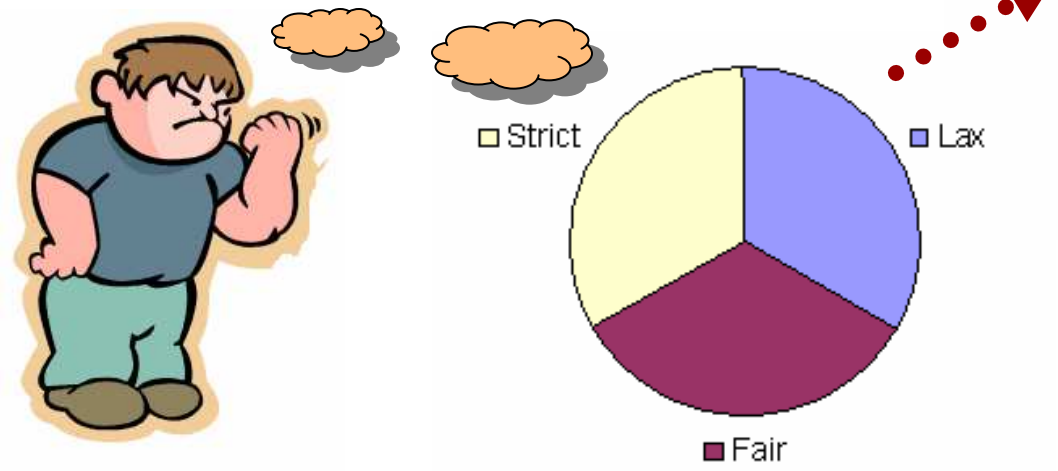
Teacher punishes bully

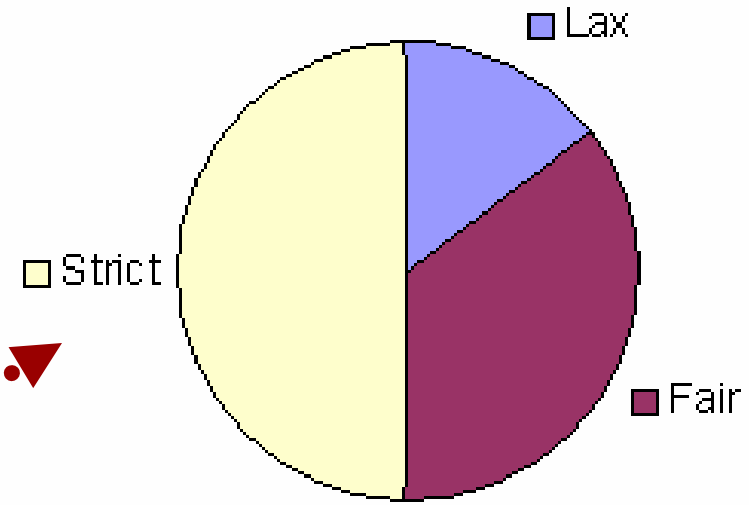
USC

ISI
Information Sciences Institute

Example: Updating Distribution

- **Based on his punishment, bully updates his probability distribution over teacher's mental models**



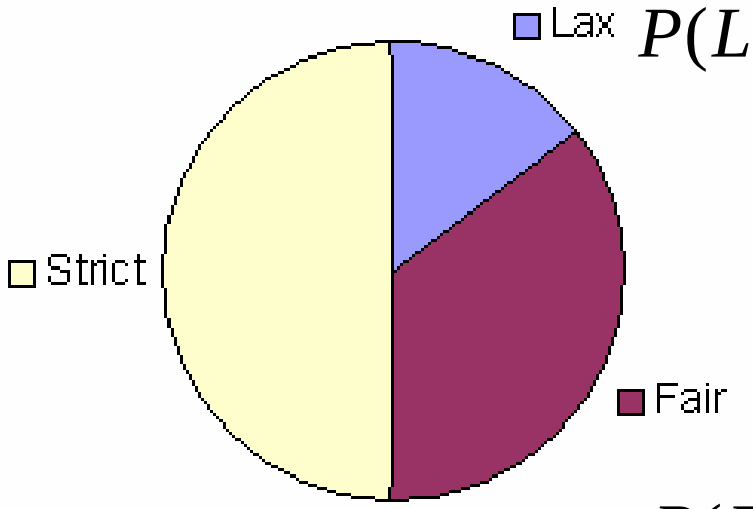


USC

ISI
Information Sciences Institute

Posterior Probabilities

$P(StrictTeacher \mid PunishBully)$



$P(LaxTeacher \mid PunishBully)$

$P(FairTeacher \mid PunishBully)$

USC

ISI
Information Sciences Institute

Calculating Posterior Probabilities

$$P(StrictTeacher \mid PunishBully) = \frac{P(StrictTeacher) \times P(PunishBully \mid StrictTeacher)}{\sum_i P(mentalModel_i) \times P(PunishBully \mid mentalModel_i)}$$

Prior Belief

Conditional Probability

?

USC

ISI Information Sciences Institute

Calculating Conditional Probability

- Conditional probability data not directly available
- However, bully can calculate teacher's expected values for a given action under different mental models

Table of Expected Values

Action	Lax	Fair	Strict
Punish Bully	.5	.75	.75
Punish Class	.4	.6	.5
Punish Observer	.3	.4	.6
Do Nothing	.8	.25	.3

USC

ISI
Information Sciences Institute

Expected Value to Conditional Probability

Bully observes teacher punishing him

$$P(PunishBully \mid StrictTeacher)$$

?

Table of Expected Values

Action	Lax	Fair	Strict
Punish Bully	.5	.75	.75
Punish Class	.4	.6	.5
Punish Observer	.3	.4	.6
Do Nothing	.8	.25	.3

USC

ISI Information Sciences Institute

Basic Assumption

- **Actions with a higher expected value should accordingly have a higher probability of being performed**

if

$$E(\mathit{punishBully}, \mathit{StrictTeacher}) > E(\mathit{doNothing}, \mathit{StrictTeacher})$$

then

$$P(\mathit{punishBully} \mid \mathit{StrictTeacher}) > P(\mathit{doNothing} \mid \mathit{StrictTeacher})$$

USC

ISI
Information Sciences Institute

Method 1: Expected Value Ratio

- **Relative expected value is good overall indicator of probability**

$$P_{ratio}(PunishBully \mid StrictTeacher) = \frac{E(PunishBully, StrictTeacher)}{\sum_i E(action_i, StrictTeacher)}$$

Table of Expected Values

Action	Lax	Fair	Strict
Punish Bully	.5	.75	.75
Punish Class	.4	.6	.5
Punish Observer	.3	.4	.6
Do Nothing	.8	.25	.3

$$\frac{.75}{.75+.5+.6+.3} = .349$$

USC

Information Sciences Institute

Ranking-Based Methods

- Relative ranking or order is good overall indicator of probability
- Convert Expected Value to Ranking

Table of Expected Values

Action	Lax	Fair	Strict
Punish Bully	.5	.75	.75
Punish Class	.4	.6	.5
Punish Observer	.3	.4	.6
Do Nothing	.8	.25	.3

Table of Rankings

Action	Lax	Fair	Strict
Punish Bully	3	4	4
Punish Class	2	3	2
Punish Observer	1	2	3
Do Nothing	4	1	1

USC

ISI
Information Sciences Institute

Linear and Exponential Ranking Methods

$$P_{rank}(PunishBully \mid StrictTeacher) = \frac{Rank(PunishBully, StrictTeacher)}{\sum_i Rank(action_i, StrictTeacher)}$$

$$P_{\exp rank}(PunishBully \mid StrictTeacher) = \frac{e^{Rank(PunishBully, StrictTeacher)}}{\sum_i e^{Rank(action_i, StrictTeacher)}}$$

Table of Rankings

Action	Lax	Fair	Strict
Punish Bully	3	4	4
Punish Class	2	3	2
Punish Observer	1	2	3
Do Nothing	4	1	1

Linear

$$\frac{4}{1+2+3+4} = .4$$

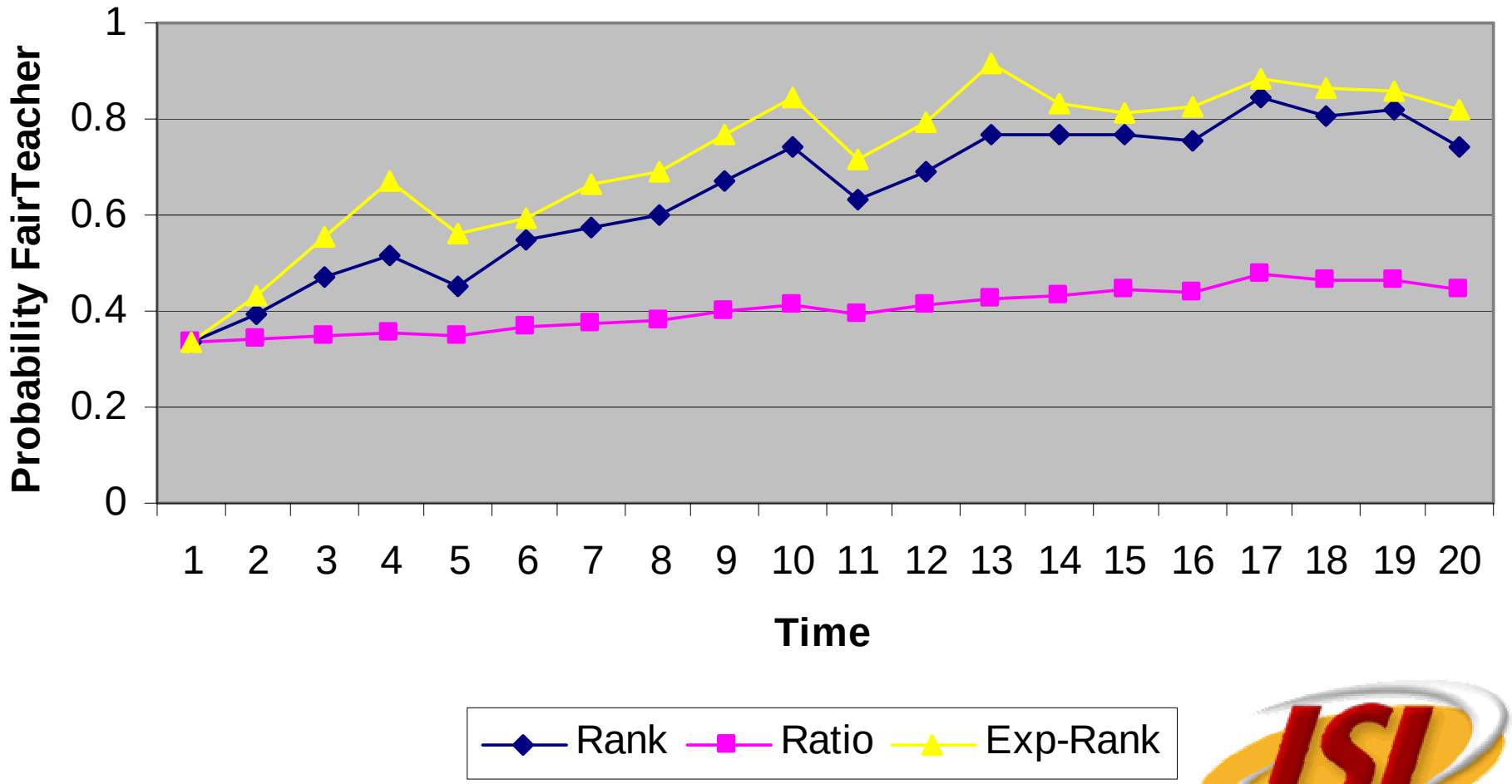
Exponential

$$\frac{e^4}{e^1+e^2+e^3+e^4} = .644$$

USC

ISI Information Sciences Institute

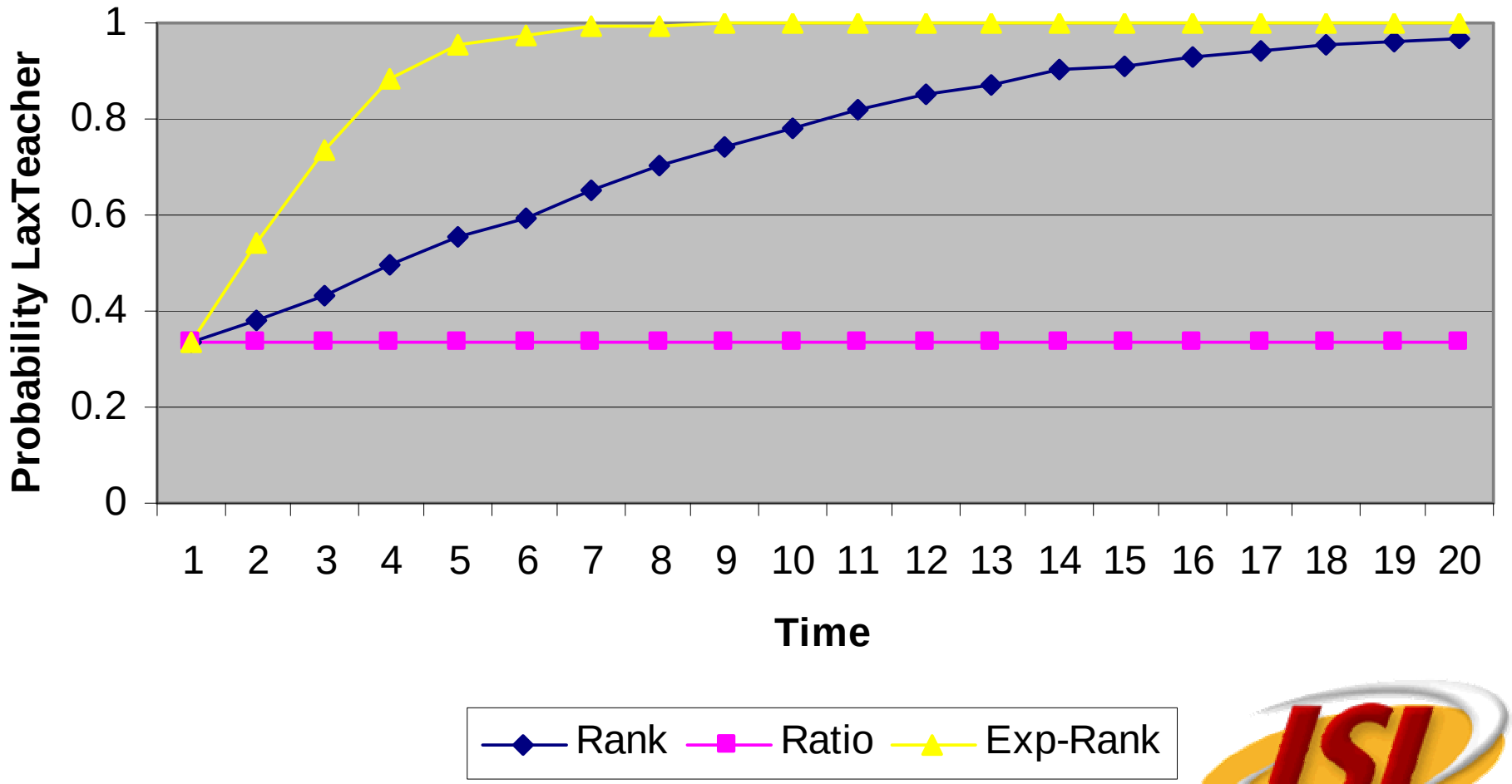
Fair Teacher



USC

ISI
Information Sciences Institute

Lax Teacher



USC

ISI
Information Sciences Institute

No Convergence in Ratio Method

- No additional preference is given for optimal actions

Expected Value Table

Observed action →

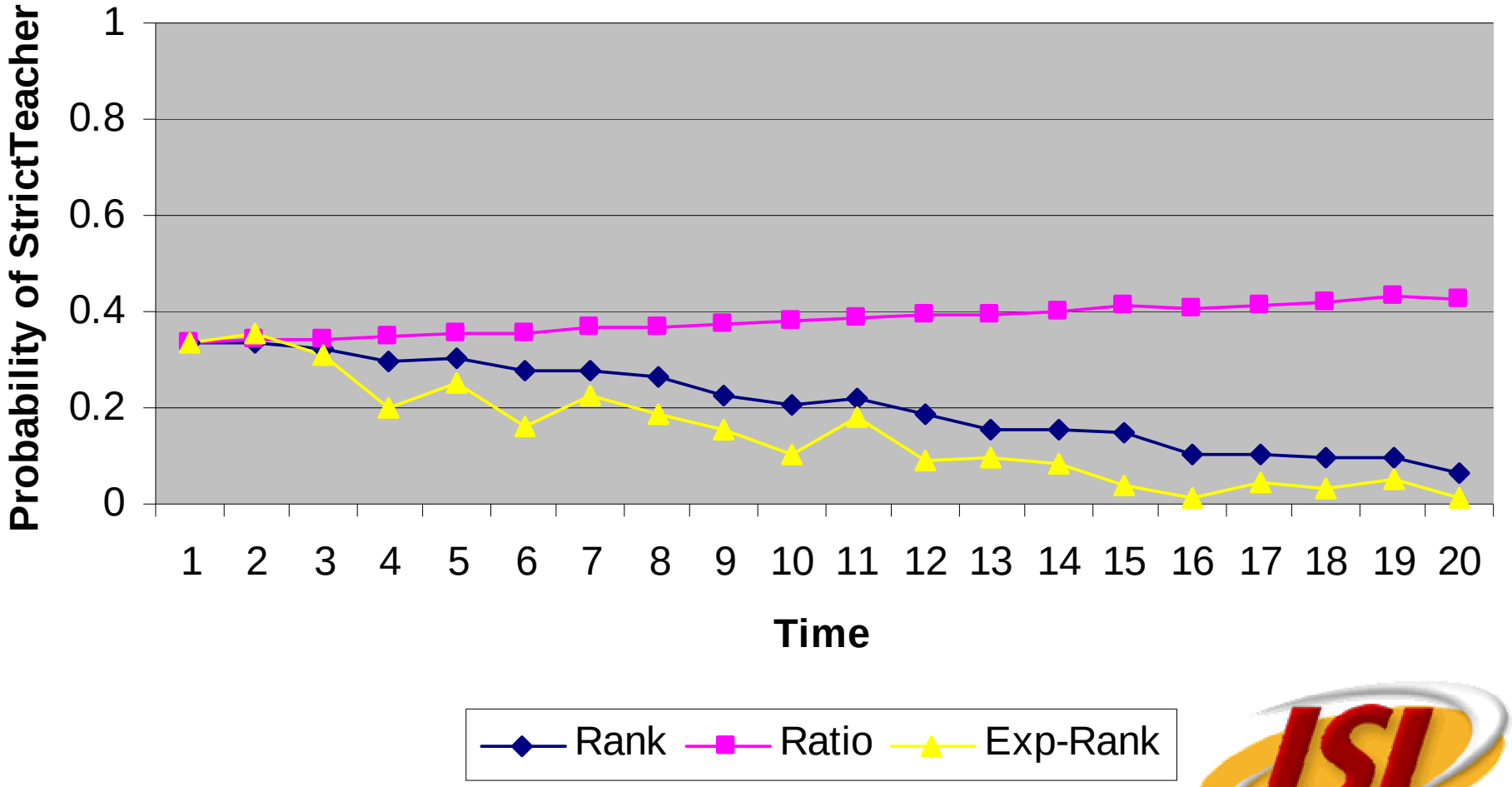
Action	Lax	Strict	Fair
Do Nothing	.9	.7	.3
Punish Class	.8	.2	.6
Punish Bully	.6	.8	.75
Punish Onlooker	.4	.4	.4

$$P_{ratio}(Nothing \mid Lax) = .33 \qquad P_{ratio}(Nothing \mid Strict) = .33$$

USC

ISI Information Sciences Institute

Strict Teacher



USC

ISI
Information Sciences Institute

What's Wrong with Ranking Methods?

- No notion of closeness

Expected Value Table

Action	Lax	Strict	Fair
Do Nothing	.9	.86	.3
Punish Class	.8	.89	.9
Punish Bully	.6	.88	.7
Punish Onlooker	.4	.87	.65

Ranking Table

Action	Lax	Strict	Fair
Do Nothing	4	1	1
Punish Class	3	4	4
Punish Bully	2	3	3
Punish Onlooker	1	2	2

USC

ISI
Information Sciences Institute

Discussion of Results

- **Ratio method**
  - Relative EV of action is accurate predictor of probability
  - Can converge slowly if EVs of actions are similar within model – no extra weight given to optimal actions
- **Ranking methods**
  - Relative ordering of actions is accurate predictor of probability
  - Much quicker convergence
  - Loses the notion of 'closeness'
- **Possible solution: Normalization across models!**

USC

ISI
Information Sciences Institute

Summary

- **Importance of mental models in constraining space**
- **Maintaining posterior probabilities over mental models**
  - Methods of calculating conditional probabilities:
    - Expected Value Ratio
    - Linear and Exponential Ranking methods
- **Preliminary experiments**
- **Identified boundary cases and issues with current methods of conditional probability calculation**

USC

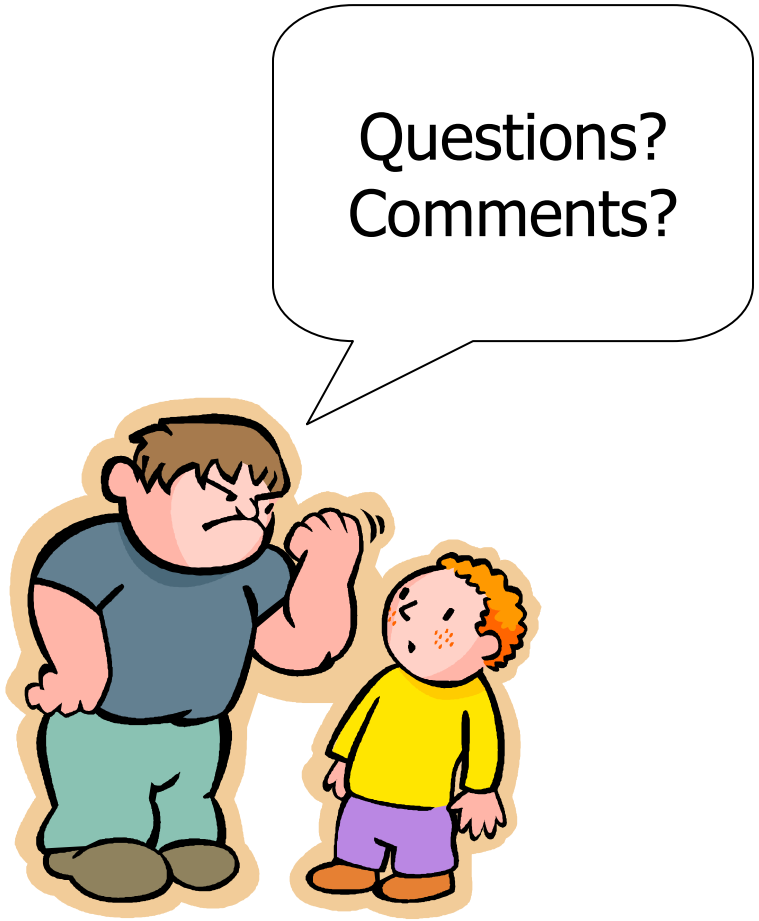
ISI
Information Sciences Institute

Future Directions

- Better methods of calculating conditional probability that deal with issues of 'closeness' and of preference of optimal actions
- More formal characterization of conditional probability calculation methods
- Imperfect memory of observations

USC

ISI
Information Sciences Institute



USC

ISI
Information Sciences Institute