Extending Fictitious Play with Pattern Recognition

Ronald Chu Gerard Vreeswijk

Department of Information and Computing Sciences
Faculty of Science, Utrecht University, The Netherlands

11th European Workshop on Multi-Agent Systems
Toulouse, 12th December 2013

Ficticious Play (1/3): The Basic Algorithm

Forecaster

- create probability distribution using observed frequencies

Responder

- Expected utility
- Best-reply correspondence

Ficticious Play (2/3): Weighted FP

Forecaster

- Muliplies observed frequencies with $0 \leq \gamma_i \leq 1$ every round

Ficticious Play (3/3): Smoothed FP

Responder

- Adds random trembles/shocks/perturbations
- Actual Utility = expected utility + benifit of exploring
  - Uses Shannon Entropy to calculate exploring utility
  - Uses Quantal Response Equilibrium to reduce chance of costly mistakes

Algorithm components we've seen so far

Forecasters

Simple uses observed distribution of play.

Weighted Adds decay to the probability distribution

Responders

Best-reply maximizes expected utility

Smoothed maximizes actual utility; adds random perturbations to the expected utility

Pattern Recognition

N-Period Ficticious Play

Forecaster

- Extends Weighted FP, uses only part of the observed history
- Create conditional probability distribution based on the last $N-1$ observations

Example: Given the history $BAAABA$ and $\gamma = 1$

- FP1: $P(x_j = A) = \frac{4}{6}$
- FP2: $P(x_j = A|A) = \frac{2}{6}$
- FP3: $P(x_j = A|BA) = \frac{1}{6}$

Cyclic pattern detection

Forecaster

- Detect simple cyclic patterns p of length $l_p \leq L$ at the end of the observed history
- Predicts $p(\text{next action of detected cycle}) = 1$ when cycle is detected

Cyclic pattern detection

Minimal conditions (weak cycle detection)

- Pattern is detected when it occures *almost* T_p times
- To make sure multiple patterns can not be detected at the same time $T_p \geq \frac{3L}{l_p}$

Necessary Conditions (strong cycle detection)

- $l_p = 1$ must appear 2 times in the last 2 rounds
- $l_p = 2$ must appear 2 times in the last 4 rounds
- $l_p > 2$ must appear *almost* 2 times in the last $2l_p - 1$ rounds

Algorithm components

Forecasters

Simple uses observed distribution of play.

Weighted Adds decay to the probability distribution

N-period WFP which uses only part of the observed history that matches the last $N-1$ rounds

Strong/Weak Detects cycles of simple patterns with length $\leq L$ at the end of observed history

Responders

Best-reply maximizes expected utility

Smoothed maximizes actual utility; adds random perturbations to the expected utility

Experiment (1/2): Algorithms

Name	**Forecast**		**Response**
FP	Simple		Best reply
SFP	Simple		Smoothed
WFP	Weighted		Best reply
SWFP	Weighted		Smoothed
FPN	N-Pattern	$N = 2, 3$	Best reply
SFPN	N-Pattern	$N = 2, 3$	Smoothed
FPwCL	Weak cycle	$L = 2, 3, 20$	Best reply
SFPwCL	Weak cycle	$L = 2, 3, 20$	Smoothed
FPsCL	Strong cycle	$L = 2, 3, 20$	Best reply
SFPsCL	Strong cycle	$L = 2, 3, 20$	Smoothed

Weight factor $\gamma = 0, 9$ and smoothing parameter $\delta = 1$

Experiment (2/2): Games

	A	B
A	$5, 5$	$0, 0$
B	$0, 0$	$3, 3$

Asymmetric Coordination Game

	A	B
A	$1, 1$	$0, 0$
B	$0, 0$	$1, 1$

Symmetric Coordination Game

	A	B
A	$3, 2$	$0, 0$
B	$0, 0$	$2, 3$

Battle of the Sexes

	A	B
A	$-1, 1$	$1, -1$
B	$1, -1$	$-1, 1$

Maching Pennies

	A	B
A	$2, 2$	$0, 3$
B	$3, 0$	$1, 1$

Prisoner's Dilemma

	A	B
A	$0, 0$	$-1, 1$
B	$1, -1$	$-10, -10$

Chicken

	A	B	C
A	$1, 0$	$0, 0$	$0, 1$
B	$0, 1$	$1, 0$	$0, 0$
C	$0, 0$	$0, 1$	$1, 0$

Shapley's Game

Results

	Prisoners	Asymmetric	Symmetric	BoS	Chicken	Pennies	Shapley	Random
FPsC – FP	=	=	=	FP	FPsC	FPsC	FPsC	FPsC
FPwC 2,3 – FP	=	=	=	FP	FPwC	FPwC	FPwC	FPwC
FPwC 20 – FP	=	=	=	≈	≈	≈	≈	≈
FPwC – FPsC	=	=	=	FPwC	FPsC	FPsC	FPsC	FPsC
FP2 – WFP	=	=	FP2	FP2	FP2	FP2	FP2	FP2
FP3 – WFP	=	=	FP3	WFP	FP3	FP3	FP3	FP3
FPN – FPC	=	=	FPC	FPC	FPN	FPN	FPN	FPN

Conclusion

N-Pattern FP

- Very effective improvement of FP
- Does not teach, only follows

Cyclic Pattern Detection

- Very effective
- Does not teach, only follows
- Higher pattern length L decreases performance of FPwC

Future Work

- Test performance against other, non-FP based, algoritms
  - Higher pattern length N for FPN
- Less static Cyclic pattern detection between the minimal and necessary conditions
- Non-FP based pattern recognition
  - Other Responders
  - No distiction between forecaster and responder