

Embedded Agency in Institutional Fields: Developing Theory from a Matching Game

Ashwin Iyengar (ashwin.iyengar15@iimb.ac.in)

Research in institutional theory is plagued with the problem of organizations being presented either as 'hypermuscular supermen' or as 'passive cultural dopes' (Suddaby, 2010). The literature on embedded agency addresses this polarity by suggesting how organizations operating in their respective institutional fields straddle this continuum. I attempt to contribute to this conversation by generating hypotheses developed from a formal model of agent-field learning through repeated interaction.

I propose to apply a simple matching game model to understand the dynamics of the interaction between embedded agents and the institutional field. This study will assume two players to be the institutional field (F), and the embedded agent (A). In each time period each of the players can make one of two decision choices 0 or 1. At the start of the game, the institutional field F choses choice 0 with probability $p_{0,F}^0$ and choice 1 with probability $1 - p_{0,F}^0$. Similarly, at the start of the game, the embedded agent A choses choice 0 with probability $p_{0,A}^0$ and choice 1 with probability $1 - p_{0,A}^0$. When the choices of the institutional field F and embedded agent A match in a time period, they each get a payoff of 1 in that period. Otherwise they each get a payoff of 0. The payoffs are captured by the matrix in Table 1

TABLE 1: Payoff Matrix

	$choice_F(t) = 0$	$choice_F(t) = 1$
$choice_A(t) = 0$	$[payoff_F(t), payoff_A(t)] = [1, 1]$	$[payoff_F(t), payoff_A(t)] = [0, 0]$
$choice_A(t) = 1$	$[payoff_F(t), payoff_A(t)] = [0, 0]$	$[payoff_F(t), payoff_A(t)] = [1, 1]$

At the end of round t , both the institutional field F and the embedded agent A update their respective probabilities $p_{t+1,F}^0$ and $p_{t+1,A}^0$ based on their prior probabilities ($p_{t,F}^0$ and $p_{t,A}^0$) and their respective learning rates ϕ_F and ϕ_A . I expect the institutional field F and embedded agent A to update their prior probabilities along the lines of the theory on reinforcement learning. I define ϕ as the learning rate function that may take any value between 0 and 1. A ϕ value of 0 indicates no learning, while a ϕ value of 1 indicates perfect learning. The institutional field F and the embedded agent A both update their prior probabilities based on their respective learning rates ϕ_F and ϕ_A respectively. The rules used to update prior probabilities are laid out in the matrix in Table 2. The intuition here is that both players win when they are aligned on the decision. The specific configuration of the alignment does not impact the payoffs beyond their being the same.

TABLE 2: Matrix of Rules for Updating Prior Probabilities

	$payoff(t) = 1$	$payoff(t) = 0$
$choice(t) = 0$	$p_{t+1}^0 = p_t^0 + \phi(1 - p_t^0)$	$p_{t+1}^0 = p_t^0 - \phi(p_t^0)$
$choice(t) = 1$	$p_{t+1}^0 = p_t^0 - \phi(p_t^0)$	$p_{t+1}^0 = p_t^0 + \phi(1 - p_t^0)$

References

Suddaby, R. 2010. Challenges for institutional theory. *Journal of Management Inquiry*, 19(1): 14–20.