Stimulus-Vector Analysis

Gail R. Berger

July 3, 2009

Stimulus Properties

Rescorla & Wagner's equations [2] also conceptualized and analyzed stimulus features and task parameters. Stimulus features, such as reward stimulus salience, of either or both US or CS, were conceptualized by α and the associative strength of conditioning contexts by V_{AX}. Stimulus features, like duration and intensity, were manifested in their λ of $\lambda - V_{AX}$ [3, p.519]. These features typically have significant impacts on task learning, irrespective of learning parameters, like β. Stimulus feature characteristics can be conceptualized as being both intrastimulus responses and paradigm characterstics.

Sutton & Barto [1, 3] noted that the above stimulus features are important contributions for mediating or interfering in reward learning. As referenced, their *theory of eligibility* included concepts like [3, p.505] stimulus features, e.g. perceived stimulus salience and contrast and other subject responses, e.g. attention, generalization, stimulus traces, etc. They noted, in addition, that learning parameters, such as interstimulus interval duration [3, p.501], or the time interval between the initial inception of CS and US occurrence, can contribute toward or impair reward task learning. The greater the distance between both, also known as delay conditioning [3, p.512], the less is the associative strength imputed onto the CS. In fact, protracted delay conditions or excessive interstimulus interval delays in their extreme, result in extinction due to unpairing of US with CS [3, p.504]. Stimulus features, such as duration, intensity, and salience apparently interact with interstimulus interval variables noted above.

Another interstimulus interval parameter referenced by Sutton & Barto is the nature of temporal pairing of both the US and CS, whether fixed or variable ratio [3, p.509]. In fixed situations, the CS consistently precedes the US and has little temporal variability between trials. In variable ratio situations,the interstimulus pairing between the two is unpredictable, variable, and without noticeable pattern. In such situations, some trials are both randomly paired, while other trials are not paired at all.

Sutton & Barto's principle reinforcement learning equations seek to conceptualize a theoretical observer's policy development, action selection, and state transitions. To account for stimulus features and task parameters Sutton & Barto [1] developed a series of equations devoted to *parameter-estimation* [1, p.564]. A stimulus *feature vector*, ϕ, represents stimulus features of salience, intensity, and duration. A task *parameter vector*, *upsilon*, delineates parameters noted above, i.e. interstimulus interval features like fixed and variable ratio characteristics. An evaluative estimate of a state, x, represents state input.

$$\upsilon_1\phi_1(x) \;+\; \cdots + \;\upsilon_n\phi_n(x) + \upsilon_{n+1} \;=\; \upsilon^T\phi(x) \quad : \; f(x) = \pi \parallel f(\upsilon) = r_{t+1}$$

Where π development is a function of the the parameter vector and associated state, $\phi(x)$ and the future return, r_{t+1}, is a function of υ.

The parameter vector, ϕ, influences an evaluative estimate x, which is the state's capability for using parameter information until the terminal (T) end of the reward task. According to Sutton & Barto (1990) $\phi(x)$ may be likened to the parameter-estimation system's internal representation of the task pattern's state, x [1, p.566]. When we look at the larger picture, we may conceptualize the following.

- The task's pattern state is a function of the parameter vector, $f(\phi) = x$. The value of the selected policy is also a function of known task parameters and the parameter vector, $f(\phi) = V^\pi(x)$.
- The future return (and motivation for completing the task) is a function of stimulus feature qualities, like reward salience, intensity, duration, and meaningfulness and feature vector, $f(\upsilon) = r_{t+1}$.
- $V_t(x_{t+1}) = \phi_t^T \upsilon_{t+1}$ or the value of the future state is composed of information derived from task parameters at a task's conclusion and parameter vector interacting with future return qualities derived from feature components and the feature vector.

In summary, Rescorla & Wagner [2] addressed stimulus qualities and task parameters in the development of their equations. With the conditioned stimulus's and context's gaining associative strength, one could assume greater predictive capability for the subject. Sutton & Barto's equations [1, 3] dissociated between reward learning and stimulus feature and task parameters. The former conceptualizations were developed over many years; the latter was conceptualized only in 1990. It is this author's belief that understanding both external and internal conditions can not only guide task-related methodology, but also can provide a better conceptualization for understanding the interaction between both external and internal conditions.

References

[1] Andrew G. Barto, Richard S. Sutton, and Christopher J. Watkins. Learning and sequential decision making. In M. Gabriel and J.W. Moore, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, pages 539–602. MIT Press, Cambridge, Massachusetts, 1990.

[2] Robert A. Rescorla and Allan R. Wagner. A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black and W.F. Prokasy, editors, *Classical Conditioning II: Current Research and Theory*, pages 64–99. Appleton-Century-Crofts, New York, New York, 1972.

[3] Richard S. Sutton and Andrew G. Barto. Time-derivative models of pavlovian reinforcement. In M. Gabriel and J.W. Moore, editors, *Learning and Computational Neuroscience: Foundations of Adapative Networks*, pages 497–537. MIT Press, Cambridge, Massachusetts, 1990.