



## 高教深耕計畫「淡江大學教師跨領域研究社群」—邀請校內教師講座(1)- 增強式學習

時間：2018-11-29 ~ 2018-12-06

地點：工學大樓 E790 會議室

增強式學習是機器學習中的一個領域，強調如何基於環境而行動，以取得最大化的預期利益。其靈感來源於心理學中的行為主義理論，即有機體如何在環境給予的獎勵或懲罰的刺激下，逐步形成對刺激的預期，產生能獲得最大利益的習慣性行為。這個方法具有普適性，因此在其他許多領域都有研究，例如博弈論、控制論、運籌學、資訊理論、仿真優化、多主體系統學習、群體智能、統計學以及遺傳算法。在運籌學和控制理論研究的語境下，增強式學習被稱作「近似動態規劃」。在最優控制理論中也有研究這個問題，雖然大部分的研究是關於最優解的存在和特性，並非是學習或者近似方面。在經濟學和博弈論中，增強式學習被用來解釋在有限理性的條件下如何出現平衡。

在機器學習問題中，環境通常被規範為馬可夫決策過程，所以許多增強式學習算法在這種情況下使用動態規劃技巧。傳統的技術和增強式學習算法的主要區別是，後者不需要關於馬可夫決策過程的知識，而且針對無法找到確切方法的大規模決策過程。

增強式學習和標準的監督式學習之間的區別在於，它並不需要出現正確的輸入/輸出對，也不需要精確校正次優化的行為。強化學習更加專注於在線規劃，需要在探索和遵從之間找到平衡。增強式學習中的「探索-遵從」的交換，在多臂老虎機問題和有限 MDP 中研究得最多。



洪智傑老師帶領社群進行研究討論。