A2C (Advantage Actor-Critic) và A3C (Asynchronous Advantage Actor-Critic) là hai phương pháp trong học tăng cường (reinforcement learning) với mục tiêu cải thiện hiệu quả của quá trình huấn luyện bằng cách kết hợp các ý tưởng từ Actor-Critic và tính toán song song. Mục tiêu của A2C A2C là một biến thể được cải thiện từ phương pháp Actor-Critic, trong đó: Actor là một mạng nơ-ron chịu trách nhiệm tạo ra các hành động cho agent, và Critic đánh giá chất lượng của hành động đó bằng cách tính Advantage – độ chênh lệch giữa giá trị hành động thực tế và kỳ vọng. Mục tiêu của A2C là huấn luyện agent một cách ổn định bằng cách tính toán các bước cập nhật liên tục cho Actor và Critic trong cùng một môi trường (đồng bộ). A2C giải quyết vấn đề không ổn định của Actor-Critic bằng cách sử dụng dữ liệu Advantage để đánh giá chính xác hơn từng hành động. Mục tiêu của A3C A3C là một cải tiến từ A2C với mục tiêu tăng tốc độ huấn luyện thông qua tính toán bất đồng bộ: Thay vì chỉ sử dụng một môi trường duy nhất, A3C sử dụng nhiều bản sao của môi trường và để các agent (hoặc worker) thực hiện độc lập trong các môi trường đó. Các agent này được huấn luyện song song và chia sẻ trọng số mạng nơ-ron với nhau, giúp cho quá trình học nhanh hơn và ổn định hơn. Mục tiêu chính của A3C là giảm thời gian huấn luyện bằng cách sử dụng các tác vụ song song để tối ưu hóa cập nhật cho Actor và Critic, đồng thời cải thiện khả năng khám phá các chiến lược khác nhau trong môi trường.