

gains instantanés (rewards) selon les pas de temps dans un épisode joué selon avec la politique apprise

