

Time Limits: Task Budgets and Interrupted Records

Pardo et al. distinguish a time limit that defines the task from a limit used to stop training interactions. For a genuinely time-limited task, remaining time affects value and should be available to the agent. Their finite return counts rewards only until the task ends; the time-aware value and update condition on the time remaining [1, Sec. 2, Eqs. 3–4].

What an Interruption Means

When the underlying task continues beyond a training timeout, they retain the successor state’s estimated value in the update. Partial-episode bootstrapping therefore separates environmental termination from an artificial reset [1, Sec. 3, Eqs. 5–6]. Its prerequisite is a successor state from which continuation can be estimated. The formula does not provide an observation or reward that was never recorded.

Figure 1 illustrates that changing how a timeout is interpreted changes the learning problem even in the same grid.

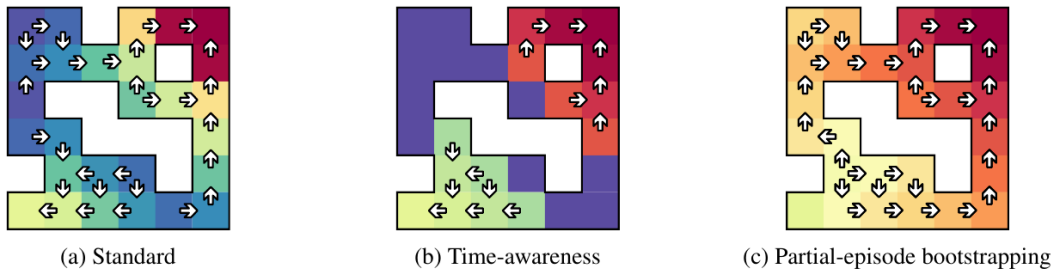


Figure 1: Pardo et al.’s toy grid: colours denote state values and arrows the learned policy under standard updates, time awareness and partial-episode bootstrapping. These are source-paper outcomes, not ARIA-NBV results. Reproduced from [1, Fig. 1].

The Three Quantities Needed for View Planning

For the thesis, distinguish the prediction horizon, the task’s remaining acquisition budget and the recorded continuation length. A completed finite-horizon return has no further rewards to count in that query. A record that ends before the requested acquisitions leaves part of that return unobserved. This is an application of the paper’s distinction, rather than a new NBV result. Its application depends on the declared task ending, not merely the length of a stored trajectory.

The paper studies time-limited and continuing control tasks. It neither decides whether an empty NBV candidate set terminates a task nor proves that fitted continuation is accurate under sparse offline action support. A missing continuation must therefore remain distinguishable from a completed finite-horizon query and a genuine task termination.

ARIA-NBV Terminology

Planning

acquisition budget – Acquisition Budget: Limit on the number of acquisitions allowed for a task.

The remaining budget counts acquisitions still available. Asking for a shorter prediction horizon does not change that budget.

return – Finite-Horizon Return: Sum of acquisition rewards over a specified prediction horizon, with discounting when specified. The core objective sums signed root-normalized rewards without discounting. A recorded return describes one realized continuation; an action value is an expected return under a specified continuation rule.

NBV – Next-Best-View: Selection of a sensor viewpoint to improve reconstruction under a limited acquisition budget. A myopic selector ranks immediate improvement; a finite-horizon planner also values views that enable useful later acquisitions.

prediction horizon – **Prediction Horizon**: Number of future acquisition rewards included in a value query. It includes the candidate acquired first and is bounded by the remaining acquisition budget; recorded trajectory length is a separate limit.

Reference

- [1] F. Pardo, A. Tavakoli, V. Levdik, and P. Kormushev, “Time Limits in Reinforcement Learning,” in *Proceedings of the 35th International Conference on Machine Learning*, in Proceedings of Machine Learning Research, vol. 80. PMLR, 2018, pp. 4045–4054. [Online]. Available: <https://proceedings.mlr.press/v80/pardo18a.html>