Reinforcement Learning Based Tractography

Tingyi Wanyan, Lantao Liu, Eleftherios Garyfallidis
Intelligent Systems Engineering, Indiana University, Bloomington

SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING
INTELLIGENT SYSTEMS ENGINEERING

Introduction

Fiber tracking is a hard problem in the area of neuroimaging [1, 2]. The goal of fiber tracking is to reconstruct pathways as close as possible to the actual anatomy. Until today researchers have introduced three main types of fiber tracking: deterministic, probabilistic or global [1]. None of these techniques have managed to radically reduce the number of invalid tracks or other major issues as reported in [1]. In this work we provide a new type of tracking based on reinforcement learning (RL) [5]. This RL-based framework allows to explore different paths by changing the positions of the start and end-points (goals) given the underlying reconstruction directions. This provides a new way of exploring the possible bundles that are available in the brain that goes beyond some of the limitations that the data provide, e.g., distinguishing crossing from diverging bundles. The method builds dynamic graphs which are updated on the fly. The nodes of the graph are shown in Fig.1.

Methods

Phase 1:Goal-Directed TD-Learning with Adaptive Expanding Graph

We extend RL by integrating a expanding-tree data structure and exploration algorithm, so that our tracker expands the tree during its exploration. The belief of the tracker is also updated and represents its knowledge learned from past experience (see Fig.1). The tree grows adaptively during the exploration, and the agent learns and updates its belief on the pathways using Temporal Difference learning(TD(0)):

$$\begin{aligned} &V(S_t) \leftarrow V(S_t) + \\ &\alpha[R_{t+1} + \gamma V(S_{t+1}) - V(S_t)], \end{aligned} \tag{1}$$

where $V(S_t)$ is state value at time t for state S_t, V_{s+1} is the estimated state value at time $t+1$, α is the learning rate, and $\gamma \in (0, 1)$ is the discounting factor for future rewards.

Phase 2: Global TD-learning

We propose to utilize a global TD-learning system where a unique copy of states, actions, and values are visible and accessible to all agents/processes. To update state values, only if a viable path is found by an agent, those states and actions along the viable path are rewarded and updated in the global system. States on failed paths are not penalized in the global system due to their possible contributions to subsequent processes.

Conclusion

Our proposed method has multiple advantages. First, the tree structure allows to explore different pathways and remember them. Since exploration paths of differing trials might overlap, the paths repeated on the same tree branch allow the tracker to make quick decisions once high reward states are encountered. Second, our method can be used to adaptively learn the space of bundles and discover the best bundle to reach a specific goal region. Finally, we are working on quantitatively showing the benefits or our approach and investigate the problem of full brain tractography using RL.

Results 1

We perform an experiment using the ISMRM 2015 tractography challenge dataset [1], we compared our RL-tracking result against deterministic tracking and the ground truth which is generated using global tracking (see [1]). The ISMRM 2015 challenge provided starting seeds positions and ending positions for each bundle, so we use the seeds positions and ending positions as the starting region and goal region as inputs to RL tracking, and use the same seeds for deterministic tracking, to make the comparison more fair we help deterministic tracking by filtering out incorrect streamlines.

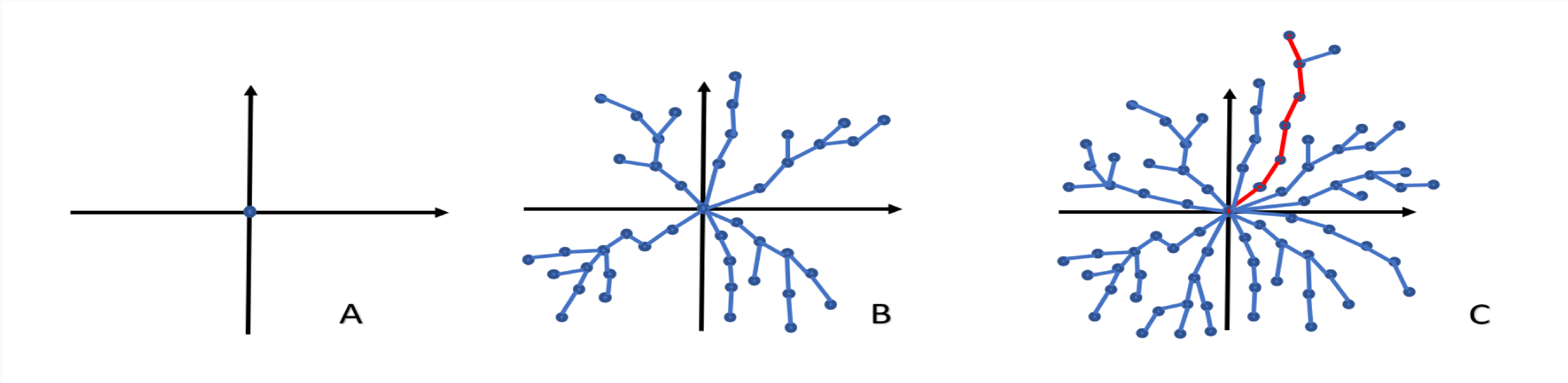

Figure 1: A.Initially, the graph system is empty. B. It generates different paths starting from the origin. Each path corresponds to a single track. C. A valid path is found and labeled red. That means a track is connecting starting region (origin) to a goal region

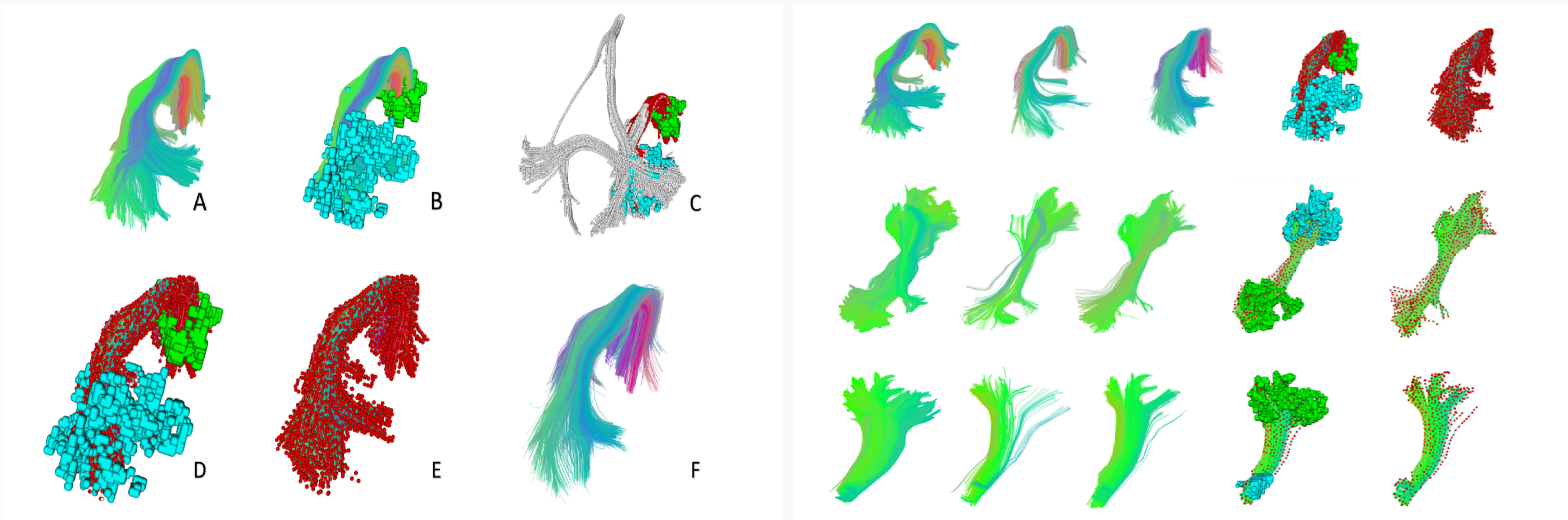

Figure 2: Left: A. Ground truth streamlines left UF bundle. B. Ground truth with starting positions and goal positions, cyan: start regions (seeds), green: goal regions. C. Result of RL tracking. Grey: nodes of expanding graph with low reward, Red: nodes with high reward, also represents the high reward global TD learning system. D. The global TD learning state nodes with starting positions and goal positions, and the corresponding streamlines found that connect the start and goal region. E. Same as D without starts/ends. F. The final extracted streamlines from RL tracking. In C-E, the graph edges are hidden for visualization purposes.; **Right**: Comparison between Deterministic Tracking and RL tracking on different bundles. Columns: GT, Det, RL, Start/End regions, graph nodes. Rows: UF, ILF and OR.

Results 2

We also test the percentage of valid tracks found against ground truth of our RL tracking algorithm comparing to deterministic tracking. See quantitative results in Table 1. Notice that we can find if a track is valid (exists in the ground truth) in a very precise way using streamline distances [4]. We conclude from these results that for most bundles, RL tracking can find more dense and valid streamlines than EuDX deterministic tracking [3].

Bundle name	Ground Truth	Det(valid)	RL(valid)	percent valid	percent not found
UF_left	5910	1138	4574	19%/77%	81%/23%
UF_right	6777	1110	1659	16.4%/24.5%	83.6%/75.5%
SLF_left	12496	3063	4283	24.5%/34.3%	75.5%/65.7%
SLF_right	11922	2128	5413	17.8%/45.4%	82.2%/54.6%
ILF_left	14777	1840	9069	12.5%/61.4%	87.5%/38.6%
ILF_right	13845	2500	7074	18.1%/51%	81.9%/49%
OR_left	8204	1082	7309	13.2%/89%	86.8%/11%
OR_right	10184	1389	1781	13.6%/17.5%	86.4%/82.5%
CST_left	7271	1625	1846	22.3%/25.4%	77.7%/74.6%
CST_right	10232	1066	2082	10.4%/20.3%	89.6%/79.7%

Table 1: Summarizing the comparisons in 10 bundles

References

[1] Maier-Hein et al., "The challenge of mapping the human connectome based on diffusion tractography", Nature Communications 2017: 8, 1349.

[2] Girard et al., "Towards quantitative connectivity analysis: reducing tractography biases, Neuroimage 2014: 98, 266-278.

[3] Garyfallidis, "Towards an accurate brain tractography", PhD thesis, University of Cambridge, 2003.

[4] Garyfallidis et al., "Quickbundles, a method for tractography simplification", Frontiers in Neuroscience 2012:6.

[5] Sutton, Barto et al., "Reinforcement learning: an introduction", MIT press, 1998.