SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Reinforcement Learning with Actor-critic for Knowledge Graph Reasoning

Linli ZHANG[1,2], Dewei LI[1,2*], Yugeng XI[1,2] & Shuai JIA[1,2]

[1]*Department of Automation, Shanghai Jiao Tong University, Shanghai* 200240*, China;*
[2]*Key Laboratory of System Control and Information Processing of Ministry of Education, Shanghai* 200240*, China*

Appendix A The whole MAP results

In the supplementary file, we give the whole tasks MAP results in Table A1 and detailed experimental result analysis. For the overall MAP shown in the last row of Table A1, our approach significantly outperforms the other three methods [1–3], which validates the strong reasoning ability of our model. For most relations, our method shows high improvement effects. However, there exist some relations that are slightly influenced due to some reason.

Since actor-critic combines policy-based and value-based networks, our model ACRL covers the original network of DeepPath. The reasoning paths found by ACRL contains almost all ones found by DeepPath. What's more, because of the evaluation of the value-based network and the interaction between 'actor' and 'critic', ACRL can find more effective reasoning paths and eliminate the wrong paths or path circles, improving the performance of fact prediction. Thus, ACRL outperforms most tasks in fact prediction substantially. However, when the original paths are already correct and complete, ACRL will show little improvement effect. Overall, ACRL shows great improvement effects than other methods.

Table A1 Fact prediction results(MAP)

Tasks	ACRL	DeepPath	TransE	TransD	Improvement
personBornInLocation	**0.4878**	0.2895	0.2652	0.0924	+68.50%
athletePlaysForTeam	**0.4384**	0.2435	0.1400	0.1494	+80.04%
teamPlaysSport	**0.4120**	0.3084	0.3567	0.1216	+33.59%
athleteHomeStadium	**0.7291**	**0.7291**	0.3810	0.0830	+0.00%
agentBelongsToOrganization	0.3287	**0.3308**	0.3498	0.1286	-0.64%
athletePlaysInLeague	**0.5199**	0.5059	0.4676	0.0762	+2.77%
personLeadsOrganization	**0.4831**	0.4785	0.3494	0.2496	+0.96%
organizationHeadQuarteredInCity	**0.6420**	0.5929	0.2569	0.1520	+8.28%
organizationHiredPerson	**0.5257**	0.4475	0.3073	0.3554	+17.47%
teamPlaysInLeague	**0.7300**	0.6532	0.7186	0.1421	+11.76%
worksFor	**0.4818**	0.4611	0.2313	0.2494	+4.49%
athletePlaysSport	0.4254	**0.5247**	0.5224	0.2651	-23.34%
Overall	**0.5170**	0.4638	0.3622	0.1720	+11.47%

Appendix B Examples of reasoning paths found by ACRL and DeepPath

Table B1 and Table B2 are some examples of reasoning paths found by ACRL and DeepPath, respectively. These relations of experimental tasks come from NELL-995 [4]. Inverses of existing relations are denoted by '_inv'.

Take for example task 'athletePlaysForTeam', DeepPath finds one more reasoning path 'athleteLedSportsTeam → teamPlaysAgainstTeam' than ACRL, which is a fake one. ACRL evaluates this policy through the critic part, and does not employ this reasoning path. Therefore, ACRL achieves high improvement in the task 'athletePlaysForTeam'. For the same reason, task 'organizationHiredPerson', 'organizationHeadQuarteredInCity' and 'worksfor' show better results by ACRL.

* Corresponding author (email: dwli@sjtu.edu.cn)

Besides, ACRL's critic part can assess actor's performance in order to prevent the emergence of the reasoning path circles, which reduces the validity of KG reasoning. In the task 'personBornInLocation', DeepPath finds some circle reasoning paths like 'personGraduatedFromUniversity → personGraduatedFromUniversity_inv → personBornInCity', 'personGraduatedFromUniversity → personGraduatedSchool_inv → personBornInCity' and 'personBelongsToOrganization → personBelongsToOrganization_inv → personBornInCity'. Conversely, ACRL eliminates these circle paths and performs better. Tasks 'athletePlaysInLeague' and 'personLeadsOrganization' have the same phenomenon.

However, the result of task 'athletePlaysSport' shows a special case that there's some probability the optimal path cannot be reserved until the end in multistep relation reasoning. It is a universality in all methods that multistep process reserves partial relations rather than the whole in each step according to the rules and computing power, which may influence the results.

In addition, ACRL shows little improvement in some tasks like 'athleteHomeStadium' and 'agentBelongsToOrganization'. According to the comparison between ACRL and DeepPath's reasoning path, it is obvious that they find almost the same reasoning paths 'athletePlaysForTeam → teamHomeStadium' and 'athleteLedSportsTeam → teamHomeStadium' about relation 'athleteHomeStadium'. ACRL can hardly improve results under the circumstance that reasoning paths are already correct and complete.

As we can see, since DeepPath fails to use the relation information entirely in the KG, it generally performs worse than our method ACRL. Our method ACRL achieves better MAP with a more compact set of learned paths. However, when there are not enough extra paths between entities, our method and DeepPath can find almost the same reasoning paths, which leads to the little improvement. In conclusion, ACRL experimental results of the fact prediction task on NELL-995 dataset shows larger performance improvement compared with the state-of-the-art methods as a whole.

Table B1 Example reasoning paths found by ACRL

Relation	Reasoning Path
personBornInLocation	personBornInCity
athletePlaysForTeam	athleteLedSportsTeam athleteHomeStadium → teamHomeStadium_inv
athleteHomeStadium	athletePlaysForTeam → teamHomeStadium athleteLedSportsTeam → teamHomeStadium
agentBelongsToOrganization	agentCollaboratesWithAgent subpartOf
athletePlaysInLeague	athletePlaysForTeam → teamPlaysInLeague athleteLedSportsTeam → teamPlaysInLeague
personLeadsOrganization	personBelongsToOrganization worksFor
organizationHeadQuarteredInCity	radioStationInCity televisionStationInCity hasOfficeInCity headQuarteredIn
organizationHiredPerson	worksFor_inv organizationTerminatedPerson personBelongsToOrganization_inv coachesTeam_inv
worksFor	organizationHiredPerson_inv personLeadsOrganization agentCollaboratesWithAgent_inv journalistWritesForPublication
athletePlaysSport	athletePlaysForTeam → teamPlaysSport athleteledSportsTeam → teamPlaysSport athletePlaysInLeague → teamPlaysInLeague_inv → teamPlaysSport athleteFlyOutToSportsTeamPosition → sportHasSportsTeamPosition_inv → sportFansInCountry → sportFansInCountry_inv

References

1 Xiong W, Hoang T, Wang W Y. DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning. In EMNLP, 2017:564-573.

2 Bordes A, Usunier N, Garcia-Duran A, et al. Translating Embeddings for Modeling Multi-relational Data. In NIPS, 2013:2787-2795.

3 Guoliang J, Shizhu H, Liheng X, et al. Knowledge graph embedding via dynamic mapping matrix. In: ACL, 2015:687–696.

4 Andrew C, Justin B, Bryan K, et al. Toward an architecture for neverending language learning. In: AAAI, 2010, 5:3

Table B2 Example reasoning paths found by DeepPath

Relation	Reasoning Path
personBornInLocation	personBornInCity
	personGraduatedFromUniversity $\rightarrow$ personGraduatedFromUniversity_inv $\rightarrow$ personBornInCity
	personGraduatedFromUniversity $\rightarrow$ personGraduatedSchool_inv $\rightarrow$ personBornInCity
	personBelongsToOrganization $\rightarrow$ personBelongsToOrganization_inv $\rightarrow$ personBornInCity
athletePlaysForTeam	athleteLedSportsTeam
	athleteHomeStadium $\rightarrow$ teamHomeStadium_inv
	athleteLedSportsTeam $\rightarrow$ teamPlaysAgainstTeam
athleteHomeStadium	athletePlaysForTeam $\rightarrow$ teamHomeStadium
	athleteLedSportsTeam $\rightarrow$ teamHomeStadium
agentBelongsToOrganization	agentCollaboratesWithAgent
	subpartOf
	agentControls_inv
	agentCollaboratesWithAgent_inv
athletePlaysInLeague	athletePlaysForTeam $\rightarrow$ teamPlaysInLeague
	athleteLedSportsTeam $\rightarrow$ teamPlaysInLeague
	athleteHomeStadium $\rightarrow$ leagueStadiums_inv
	athletePlaysSport $\rightarrow$ teamPlaysSport_inv $\rightarrow$ teamPlaysInLeague
	athletePlaysForTeam $\rightarrow$ teamPlaysAgainstTeam_inv $\rightarrow$ teamPlaysInLeague
	athletePlaysSport $\rightarrow$ teamPlaysSport_inv $\rightarrow$ teamPlaysAgainstTeam_inv $\rightarrow$ teamPlaysInLeague
	athleteFlyOutToSportsTeamPosition $\rightarrow$ athleteFlyOutToSportsTeamPosition_inv $\rightarrow$ athletePlaysSport $\rightarrow$ teamPlaysSport_inv $\rightarrow$ teamPlaysInLeague
personLeadsOrganization	personBelongsToOrganization
	worksFor
	organizationTerminatedPerson_inv
	mutualProxyFor_inv
	organizationHiredPerson_inv
	agentCollaboratesWithAgent_inv
	worksFor $\rightarrow$ worksFor_inv $\rightarrow$ worksFor
organizationHeadQuarteredInCity	radioStationInCity
	televisionStationInCity
	hasOfficeInCity
	headQuarteredIn
	competesWith_inv $\rightarrow$ headQuarteredIn
organizationHiredPerson	worksFor_inv
	organizationTerminatedPerson
	personBelongsToOrganization_inv
	coachesTeam_inv
	mutualProxyFor
	personLeadsOrganization_inv
worksFor	organizationHiredPerson_inv
	personLeadsOrganization
	agentCollaboratesWithAgent_inv
	journalistWritesForPublication
	topmemberOfOrganization
athletePlaysSport	athletePlaysForTeam $\rightarrow$ teamPlaysSport
	athleteledSportsTeam $\rightarrow$ teamPlaysSport
	athletePlaysInLeague $\rightarrow$ teamPlaysInLeague_inv $\rightarrow$ teamPlaysSport
	athletePlaysInLeague $\rightarrow$ leagueStadiums $\rightarrow$ sportUsesStadium_inv
	athletePlaysInLeague $\rightarrow$ subpartOfOrganization_inv $\rightarrow$ teamPlaysSport