KSG: Knowledge and Skill Graph

Feng Zhao
Westlake University
Hangzhou, Zhejiang, China
zhaofeng@westlake.edu.cn

Ziqi Zhang
Westlake University
Hangzhou, Zhejiang, China
zhangzq20@mails.tsinghua.edu.cn

Donglin Wang*
Westlake University
Westlake Institute for Advanced Study
Hangzhou, Zhejiang, China
wangdonglin@westlake.edu.cn

ABSTRACT

The knowledge graph (KG) is an essential form of knowledge representation that has grown in prominence in recent years. Because it concentrates on nominal entities and their relationships, traditional knowledge graphs are static and encyclopedic in nature. On this basis, event knowledge graph (Event KG) models the temporal and spatial dynamics by text processing to facilitate downstream applications, such as question-answering, recommendation and intelligent search. Existing KG research, on the other hand, mostly focuses on text processing and static facts, ignoring the vast quantity of dynamic behavioral information included in photos, movies, and pre-trained neural networks. In addition, no effort has been done to include behavioral intelligence information into the knowledge graph for deep reinforcement learning (DRL) and robot learning. In this paper, we propose a novel dynamic knowledge and skill graph (KSG), and then we develop a basic and specific KSG based on CN-DBpedia. The nodes are divided into entity and attribute nodes, with entity nodes containing the agent, environment, and skill (DRL policy or policy representation), and attribute nodes containing the entity description, pre-train network, and offline dataset. KSG can search for different agents' skills in various environments and provide transferable information for acquiring new skills. This is the first study that we are aware of that looks into dynamic KSG for skill retrieval and learning. Extensive experimental results on new skill learning show that KSG boosts new skill learning efficiency.

CCS CONCEPTS

• **Information systems → Information retrieval diversity**.

KEYWORDS

Knowledge and Skill Graph, Skill Retrieval, Knowledge Graph

1 INTRODUCTION

Knowledge Graph (KG), first announced by Google in 2012, is a popular and efficient model for knowledge representation that has attracted the interest of researchers in related fields [2, 27, 31]. A knowledge graph is a knowledge base of information about entities (e.g., people and organizations) that represent the many entities and their relationships in a domain using a collection of subject-predicate-object triplets. Each triplet is also referred to as a fact. Nodes represent entities, while edges reflect relationships between things in a knowledge graph [1, 14, 16, 17, 23, 25, 35]. Nowadays, along with the continuous development of intelligent information service applications, the knowledge graph has been widely applied to many fields, such as question-answering, recommendation, text generation and so on [13, 15, 22, 40–42].

However, there is a wealth of event information available in the world, such as the most recent news stories, which conveys dynamic and procedural knowledge. As a result, event-centric knowledge representation forms such as Event KG (EKG) have been a popular research topic [6, 7, 26, 32, 36]. To address the needs of rapid retrieval and concise representation of event-related information, EKG can filter and structure information about events reported in texts [10, 18, 37–39]. Meanwhile, in order to capture dynamic information of events, EKG considers action, participant, time, and location to extract event-event relations including temporal and causal relations [9, 11, 24]. However, existing methods on KG or EKG usually focus on text processing and disregard dynamic behavior information, making it impossible to search for and understand the behavior or skills of humans and agents [19–21, 28, 29].

In this paper, we present a novel concept of Knowledge and Skill Graph (KSG) to address above problems. Based on CN-DBpedia [34], we add additional nodes and relations between nodes to the knowledge graph. KSG has two types of nodes, entities and attributes, as well as two types of directed edges, which indicate entity-entity and entity-attribute relationships. Specifically, entity nodes include fact nodes (such as person, etc.), environment nodes, and skill nodes. Entity description, skill display, pre-train network, and offline dataset are examples of attribute nodes. KSG retains initial information retrieval and question-and-answer functionalities. Meanwhile, KSG can search for and display the skills of different agents in various environments, as well as give transferable knowledge for learning new skills. If we want to train a new skill, we can search for learned skills in KSG related to the new skill as a pre-training networks and use them to train the new skill. Our experiments demonstrate that KSG can achieve effective skill searching and learning. The main contributions of this paper can be summarized as follows:
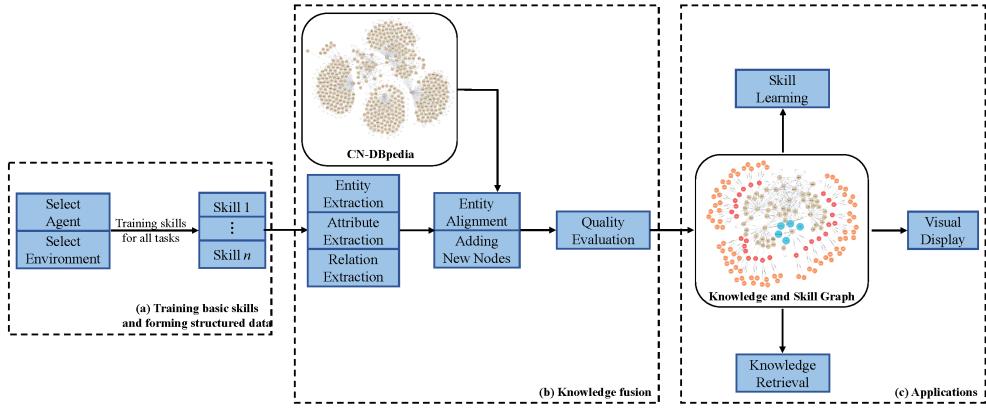
Figure 1: The architecture of knowledge and skill graph. We first train basic skills and form structured data in (a). Then, we extract dynamic behavioral information from basic skills and construct a preliminary but specific KSG based on CN-DBpedia in (b). Finally, we apply KSG to achieve knowledge retrieval, visual display and skill learning in (c).

- In order to extend KG to deal with dynamic behaviors, we propose a novel concept of KSG to simultaneously process static and dynamic knowledge.
- We propose and visualize a preliminary and specific KSG. To the best of our knowledge, our work is the first to implement skill retrieval and skill reuse by introducing behavioral intelligence into KG.
- Extensive experiments demonstrate the effectiveness of KSG in a variety of tasks, such as QA, knowledge retrieval and skill learning.

2 KNOWLEDGE AND SKILL GRAPH (KSG)

Our aim is to establish a knowledge and skill graph (KSG) that not only retains the original static facts of the knowledge graph but also deals with dynamic behavior information. As a result, KSG can be utilized for both text-based Q&A systems, retrieval and recommendation, and skill retrieval, as well as providing transferable knowledge for learning new skills. In this paper, we emphasize the last functionality. As shown in Figure 1, knowledge and skill graph is constructed based on CN-DBpedia and the construct process is divided into three parts: a) Training Skills for Data Preparation; b) knowledge Fusion; c) Applications. We first introduce the methods and process of training basic skills and the storage of skills. Then, we elaborate on the details of knowledge fusion. Finally, we explain the main applications and the details are described in section 3.

2.1 Training Skills for Data Preparation

In the processing of constructing KSG, we firstly train a great deal of basic skills of different agents in different environments. In this paper, the agent Humanoid, Ant and Half Cheetah are from Mujoco [30]. Environment plane is smooth ground and obstacle represents an obstructed ground. Then, we design different tasks according to the direction of walk. Finally, we consider the 18-DoF quadruped robot. For simulative quadruped robot, we design three environments: plane, stair, and irregular plane, and only consider the walk task. Specifically, we also store the walk skill of real quadruped robot in the environment plane. In this paper, we use SAC [12] to train

basic skills. By choosing different tasks, agents and environments, we have trained 23 different skills. For each skill, we store such trained network and offline data as knowledge which will be used to construct KSG.

2.2 Knowledge Fusion

In this paper, our preliminary KSG is based on the CN-DBpedia dataset. During the process of the KSG construction, we use neo4j and py2neo to import the triple relationship of CN-DBpedia. Among them, Neo4j is a powerful graph database tool that can search specific nodes with O(1) time complexity, and Py2neo is a python package that allows us to manipulate the neo4j database in a python environment. Once we have imported the CN-Dbpedia database, we can directly build the KSG based on it by adding a variety of skills, environments and agents. Different from traditional KG, we innovatively introduce dynamic behavior information into KG and add new nodes where added entity nodes consist of agents, environments and skills, as well as attribute nodes include entity description, skill display, pre-trained network and offline dataset.

To obtain new nodes about skills, we first need knowledge extraction from trained basic skills. In this paper, the knowledge extraction is divided into three types: entity extraction, attribute extraction and relation extraction. For entity extraction, we use named-entity recognition (NER) to classify entity into pre-defined categories such as human, plane, walk up, walk right and so on. In this part, we extract new entity nodes including agent, environment and skill. After entity extraction, the attribute extraction is to define the attribute or description of entity. For relation extraction, we use rule-based and dictionary-based methods to extract the relationships among the entities and attributes. Relation extraction is to find the relations between entity-entity and entity-attribute. For example, *"huanmoid"* is one of the most popular agents which have been broadly used in various Reinforcement learning researches, and has many human's features, so we use NER to find entity node *"human"* in CN-Dbpedia, and then add some pre-defined skill entity nodes (such as *"Walk_Up"*), environment entity nodes (*"Plane"*, *"Obstacle"*) to it.
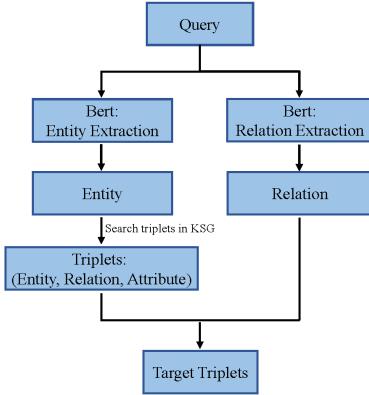
Figure 2: Process of KSGQA

2.3 Applications

Except for the construction of KSG, we also build the KSG Question-Answer system (KSGQA) based on the KSG to facilitate the use of prior knowledge (in this paper, prior knowledge means the trained models as well as identity representations). This system can interpret and encode the specified queries in order to acquire target triplets. As shown in Figure 2, our KSGQA utilizes fine-tuned Bert [8] to encode query. In this way, we can identify entity and relations between entities from the query. According to the entity, we can first obtain the corresponding triplets including entity, relation and attribute from KSG, and then we use the relation to determine the target triplets. For example, when inputting a query *"Do you know what skill humans have?"*, we can use Bert to extract entity *"human"* and *"skill"*. Meanwhile, the relation *"have"* is used to obtain target triplets such as *"{'human', 'have', 'walk_left'}"*, *"{'human', 'have', 'walk_right'}"* and so on. Based on the KSGQA, we can realize some functionalities such as knowledge retrieval, visual display and skill learning.

For knowledge retrieval, we can retrieve those skills that an agent has learned and those data to be called. After we obtain a list of the agent's skills, we can display those skills as well. For example, when we input *"Do you have know what skills {agent_name} have?"* and obtain a list of the agent's skills, we can input query *"Can you show {agent_name} {skill_name} in the {environment_name}?"* to load the video and show such skill. In addition, we store corresponding trained model into KSG so that we just have to query this KSGQA system to load the trained model. For example, when we input *"Can you search {agent_name} {skill_name} in the {environment_name}?"*, we can get the trained model and offline data of the agent. Thus, if we need a new skill that is not available in KSG, the trained model and offline data can be regarded as pre-trained model and training data to learn new skill [3–5, 33]. In this paper, we define two selection strategies for pre-training models based on environmental differences and task differences. When performing the same task in different environments, we choose the skill model with the highest similarity as the pre-training model according to the environment similarity which is determined by calculating the Euclidean distance between the sampled states in different environments. When

performing different tasks in the same environment, we select the pre-training model according to the task similarity.

3 EXPERIMENTS AND APPLICATIONS OF KSG

We aim to establish a KSG which can be used to search skills and provide transferable knowledge for learning new skills. In this section, we first show KSG's Q&A system and skill retrieval capabilities. Then, we use KSG to retrieve and provide related skill model and offline data for learning new skills.

3.1 Knowledge retrieval and display

In this paper, we construct a preliminary and specific KSG based on CN-DBpedia. Compared with traditional knowledge graph, KSG focuses on behavioral intelligence. We retain the original functions of the knowledge graph, meanwhile introducing new dynamic knowledge and skills into it. Therefore, we also design a corresponding knowledge and skill graph question answer system. The important functions of KSGQA are knowledge retrieval and display. As shown in Figure 3, we show the basic functions of knowledge retrieval. We can see that KSG can be used to retrieval existing skills as shown in Figure 3 (a), and then when we need the networks and offline data of these skills, we can call and download them from KSG in Figure 3 (b). In addition, KSG can display the stored skills. For example, when we input *"Can you show ant walking down in the plane?"*, the KSGQA will load the video and show how the ant walks toward the down. Moreover, if we need a new skill which is not existing in KSG, KSG is able to quickly select most related base skills as pre-trained models to help learn new skills.

(a) Skill retrieval

(b) Pre-trained model retrieval

Figure 3: KSGQA for knowledge retrieval.

3.2 Skill Storage and New Skill Learning

In this paper, we consider agents from Mujoco and real quadruped robot as shown in Figure 4. Each agent has different skills in different environments. We add skill nodes and corresponding attribute nodes to store these skills in KSG, where the stored knowledge includes pre-trained model, video and offline data.
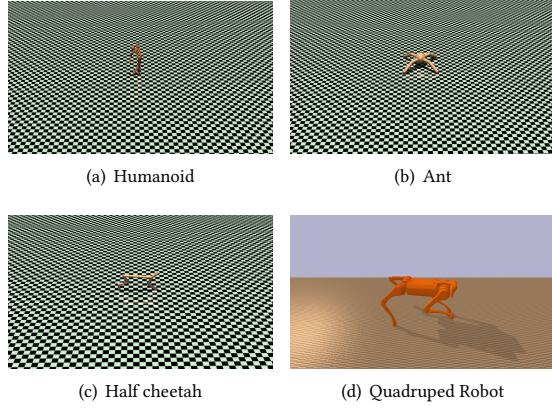
(a) Humanoid　　　　　　　　(b) Ant

(c) Half cheetah　　　　　　　(d) Quadruped Robot

Figure 4: All agents being considered in our specific KSG.

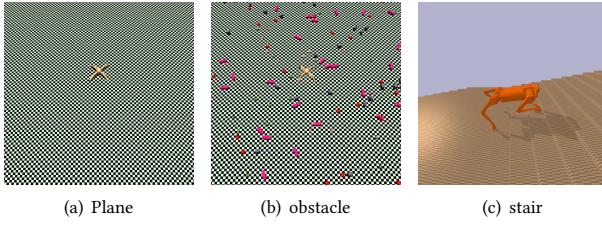(a) Plane　　　　　(b) obstacle　　　　　(c) stair

Figure 5: All three environments in our preliminary but specific KSG including plane, obstacle and stair.

In order to enrich our KSG, we design different environments to complete each task. A part of these environments are shown in Figure 5, including *Plane*, *obstacle*, and *stair*. By performing different tasks in different environments, we can acquire different skills for each agent. These stored skills are transferable knowledge that can be used to learn new skills. In this part, we use KSG to help agent learn new skills in different environment. Actually, the transferable knowledge includes pre-trained neural network and offline dataset. Therefore, we can directly use one of the most relevant models as a pre-training model for new skills, or we can use multiple related skills to combine and learn a new skill.
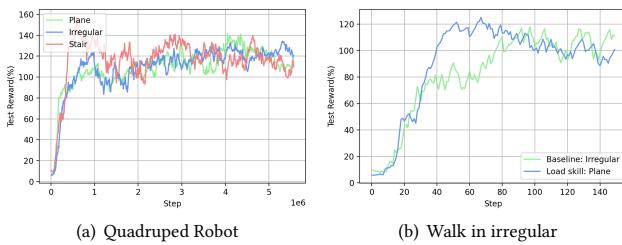

(a) Quadruped Robot　　　　　(b) Walk in irregular

Figure 6: Test reward of Quadruped Robot and learn new skills in new environment using pre-trained model.

In this paper, we can load stored skill model from KSG as pre-train model to learn new skill. In Figure 6 (a), we show the stored skill model's test reward of real Quadruped Robot. If we now need to acquire the skills of walk for quadruped robot in environment irregular, we can select the most relevant skills from KSG as a pre-training model. This problem belongs to performing the same task in different environments. We first calculate the similarity of environment between irregular, plane and stair. In the above question, we select the skills with the highest similarity by calculating task similarity (Walk in environment plane) as the pre-training model to learn new skill walk in environment irregular. As show in Figure 6 (b), we can see that loading related pre-training models can improve training efficiency and reduce nearly half of the training time compared with direct training from scratch.
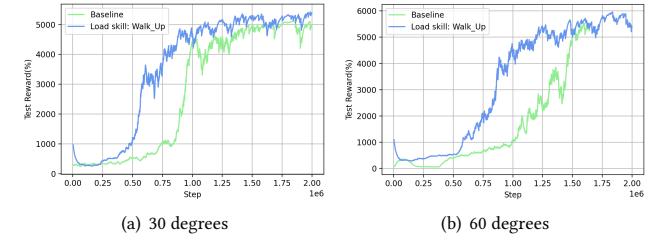

(a) 30 degrees　　　　　　　(b) 60 degrees

Figure 7: Obliquely upward walk at a different Angle from the horizontal direction

On the other hand, if we need a new skill which is performing different tasks in the same environment, we can select the pre-training model according to the task similarity. When we need two skills that are obliquely upward walk at a different angle from the horizontal direction (30 degrees, 60 degrees), we can automatically select skill walk right and walk up as pre-training models according to the task similarity. As shown in figure 7, we can observe that loading pre-training skills according to task similarity can also improve the learning efficiency of new skills.

4 CONCLUSION AND FUTURE WORK

In this paper, we propose a novel concept of KSG to simultaneously process static and dynamic knowledge. KSG introduces dynamic behavioral information to KG, which can also implement skill retrieval and skill reuse. Existing methods on KG usually focus on text pro-cessing and static fact, so we extend KG to deal with dynamic behaviors. KSG retains the original function of the knowledge graph and adds new functions including skill retrieval, visual display of skill and knowledge transferring for new skill learning. In experiments, we show all of KSG's capabilities in detail. It indicates that constructing KSG is significant and valuable for storing and utilizing dynamic behavioral information.

Although our KSG now is preliminary, it is very meaningful. In the future, we will continue to expand and improve KSG for real applications. On the one hand, KSG will provide a large amount of basic skills and offline data for reinforcement learning, meta learning, imitation learning and so on. On the other hand, KSG will also provide more complex relation between skills, agents and environments for skill learning and reasoning.

REFERENCES

[1] Bilal Abu-Salih, Marwan Al-Tawil, Ibrahim Aljarah, Hossam Faris, Pornpit Wongthongtham, Kit Yan Chan, and Amin Beheshti. 2021. Relational learning analysis of social politics using knowledge graph embedding. *Data Mining and Knowledge Discovery* (2021), 1–40.

[2] Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. 2009. Dbpedia-a crystallization point for the web of data. *Journal of web semantics* 7, 3 (2009), 154–165.

[3] Zhengyu Chen, Jixie Ge, Heshen Zhan, Siteng Huang, and Donglin Wang. 2021. Pareto Self-Supervised Training for Few-Shot Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 13663–13672.

[4] Zhengyu Chen and Donglin Wang. 2021. Multi-Initialization Meta-Learning with Domain Adaptation. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1390–1394.

[5] Zhengyu Chen, Ziqing Xu, and Donglin Wang. 2021. Deep transfer tensor decomposition with orthogonal constraint for recommender systems. In *The Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI*, Vol. 2021. 3.

[6] Anthony Colas, Ali Sadeghian, Yue Wang, and Daisy Zhe Wang. 2021. Event-Narrative: A large-scale Event-centric Dataset for Knowledge Graph to Text Generation.

[7] Tarcísio Souza Costa, Simon Gottschalk, and Elena Demidova. 2020. Event-QA: A Dataset for Event-Centric Question Answering over Knowledge Graphs. *arXiv: Computation and Language* (2020).

[8] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[9] Xiao Ding, Zhongyang Li, Ting Liu, and Kuo Liao. 2019. ELG: An Event Logic Graph. *arXiv: Artificial Intelligence* (2019).

[10] Goran Glava and Jan najder. 2015. Construction and evaluation of event graphs. *Natural Language Engineering* 21 (2015), 607–652.

[11] Simon Gottschalk and Elena Demidova. 2018. EventKG: A Multilingual Event-Centric Temporal Knowledge Graph. In *European Semantic Web Conference*.

[12] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*. PMLR, 1861–1870.

[13] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. 2015. DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Social Work* 6 (2015), 167–195.

[14] Chen Li, Xutan Peng, Yuhang Niu, Shanghang Zhang, Hao Peng, Chuan Zhou, and Jianxin Li. 2021. Learning graph attention-aware knowledge graph embedding. *Neurocomputing* 461 (2021), 516–529.

[15] Linfeng Li, Peng Wang, Jun Yan, Yao Wang, Simin Li, Jinpeng Jiang, Zhe Sun, Buzhou Tang, Tsung-Hui Chang, Shenghui Wang, et al. 2020. Real-world data medical knowledge graph: construction and applications. *Artificial intelligence in medicine* 103 (2020), 101817.

[16] Zelong Li, Jianchao Ji, Zuohui Fu, Yingqiang Ge, Shuyuan Xu, Chong Chen, and Yongfeng Zhang. 2021. Efficient Non-Sampling Knowledge Graph Embedding. In *Proceedings of the Web Conference 2021*. 1727–1736.

[17] Zhifei Li, Hai Liu, Zhaoli Zhang, Tingting Liu, and Neal N Xiong. 2021. Learning knowledge graph embedding with heterogeneous relation attention networks. *IEEE Transactions on Neural Networks and Learning Systems* (2021).

[18] Zhongyang Li, Sendong Zhao, Xiao Ding, and Ting Liu. 2017. EEG: Knowledge Base for Event Evolutionary Principles and Patterns.

[19] Jinxin Liu, Hao Shen, Donglin Wang, Yachen Kang, and Qiangxing Tian. 2021. Unsupervised Domain Adaptation with Dynamics-Aware Rewards in Reinforcement Learning. *Advances in Neural Information Processing Systems* 34 (2021), 28784–28797.

[20] Jinxin Liu, Donglin Wang, Qiangxing Tian, and Zhengyu Chen. 2022. Learn goal-conditioned policy with intrinsic motivation for deep reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 7558–7566.

[21] Jinxin Liu, Hongyin Zhang, and Donglin Wang. 2022. DARA: Dynamics-Aware Reward Augmentation in Offline Reinforcement Learning. *arXiv preprint arXiv:2203.06662* (2022).

[22] Haithem Mezni. 2021. Temporal Knowledge Graph Embedding for Effective Service Recommendation. *IEEE Transactions on Services Computing* (2021).

[23] Haithem Mezni, Djamal Benslimane, and Ladjel Bellatreche. 2021. Context-aware service recommendation based on knowledge graph embedding. *IEEE Transactions on Knowledge and Data Engineering* (2021).

[24] Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard. 2016. Building event-centric knowledge graphs from news. *Journal of Web Semantics* 37 (2016), 132–151.

[25] Andrea Rossi, Denilson Barbosa, Donatella Firmani, Antonio Matinata, and Paolo Merialdo. 2021. Knowledge graph embedding for link prediction: A comparative analysis. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 15, 2 (2021), 1–49.

[26] Charlotte Rudnik, Thibault Ehrhart, Olivier Ferret, Denis Teyssou, Raphaël Troncy, and Xavier Tannier. 2019. Searching News Articles Using an Event Knowledge Graph Leveraged by Wikidata. *arXiv: Computation and Language* (2019).

[27] Fabian M Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web*. 697–706.

[28] Qiangxing Tian, Jinxin Liu, Guanchu Wang, and Donglin Wang. 2021. Unsupervised discovery of transitional skills for deep reinforcement learning. In *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.

[29] Qiangxing Tian, Guanchu Wang, Jinxin Liu, Donglin Wang, and Yachen Kang. 2021. Independent skill transfer for deep reinforcement learning. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*. 2901–2907.

[30] Emanuel Todorov, Tom Erez, and Yuval Tassa. 2012. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 5026–5033.

[31] Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a free collaborative knowledgebase. *Commun. ACM* 57, 10 (2014), 78–85.

[32] Jie Wu, Xinning Zhu, Chunhong Zhang, and Zheng Hu. 2020. Event-centric Tourism Knowledge Graph—A Case Study of Hainan. In *Knowledge Science, Engineering and Management*.

[33] Teng Xiao, Zhengyu Chen, Donglin Wang, and Suhang Wang. 2021. Learning How to Propagate Messages in Graph Neural Networks. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 1894–1903.

[34] Bo Xu, Yong Xu, Jiaqing Liang, Chenhao Xie, Bin Liang, Wanyun Cui, and Yanghua Xiao. 2017. CN-DBpedia: A Never-Ending Chinese Knowledge Extraction System. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 428–438.

[35] Yingying Xue, Jiahui Jin, Aibo Song, Yingxue Zhang, Yangyang Liu, and Kaixuan Wang. 2021. Relation-based multi-type aware knowledge graph embedding. *Neurocomputing* 456 (2021), 11–22.

[36] Chengbiao Yang, Weizhuo Li, Xiaoping Zhang, Runshun Zhang, and Guilin Qi. 2019. A Temporal Semantic Search System for Traditional Chinese Medicine Based on Temporal Knowledge Graphs. In *International Semantic Technology Conference*.

[37] Min Zhang, Siteng Huang, Wenbin Li, and Donglin Wang. 2022. Tree Structure-Aware Few-Shot Image Classification via Hierarchical Aggregation. In *Proceeding of the 17th European Conference on Computer Vision, ECCV*.

[38] Min Zhang, Siteng Huang, and Donglin Wang. 2022. Domain Generalized Few-Shot Image Classification via Meta Regularization Network. In *ICASSP*. 3748–3752.

[39] Min Zhang, Donglin Wang, and Sibo Gai. 2020. Knowledge Distillation for Model-Agnostic Meta-Learning. In *European Conference on Artificial Intelligence, ECAI*.

[40] Feng Zhao, Tiancheng Huang, and Donglin Wang. 2022. Graph Few-Shot Learning via Restructuring Task Graph. *IEEE Transactions on Neural Networks and Learning Systems* (2022).

[41] Feng Zhao and Donglin Wang. 2021. Multimodal Graph Meta Contrastive Learning. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. 3657–3661.

[42] Feng Zhao, Donglin Wang, and Xintao Xiang. 2021. Multi-Initialization Graph Meta-Learning for Node Classification. In *Proceedings of the 2021 International Conference on Multimedia Retrieval*. 402–410.