Representing Goals as Guidance

Penglin Cai
Yuanpei College
Peking University
cpl@stu.pku.edu.cn

Abstract

Representing goals has always been an important work in machine learning and agent systems. However, many end-to-end trained agents usually represents goals from a certain kind of perception, constructing the correspondence between sensory organs of humans and goal representations. This essay summarizes several commonly used methods of goal representation, from the perspective of vision, language, and auditory sense. In addition, some other previous work simply uses implicit embeddings or latent variables to represent goals, about which we will also discuss. As a combination of the aforementioned content, some recent work uses multi-modal representations of goals to convey much more information, which is worth considering for future researches.

1 Introduction

Humans do have goals, and goals can be divided into a great many categories. Short-term goals can serve as intentions, such as picking up an object or walking to somewhere. Long-term goals can serve as ideals, driving people to strive forward.

Ever since young, humans begin to understand goals. Experimental results have shown that 6-month-old infants can already infer others' intentions from their actions [22]. The 12-month-old have recognized that speech can communicate unobservable intentions [33].

In general, goals play a role in guiding the aim and direction of actions and behaviors, thus influencing people's decisions. Inspired by this, researchers have developed a complete set of theories of goal-as-conditions, modeling decision processes such as Markov Decision Process (MDP). In recent studies, goals are widely modeled in Goal-Conditioned Reinforcement Learning (GCRL) (Figure 1), which can be used for multi-agent systems as well as robot manipulation. Given the importance of goals in these systems, how to represent goals for agents to learn has been a significant problem in machine learning.

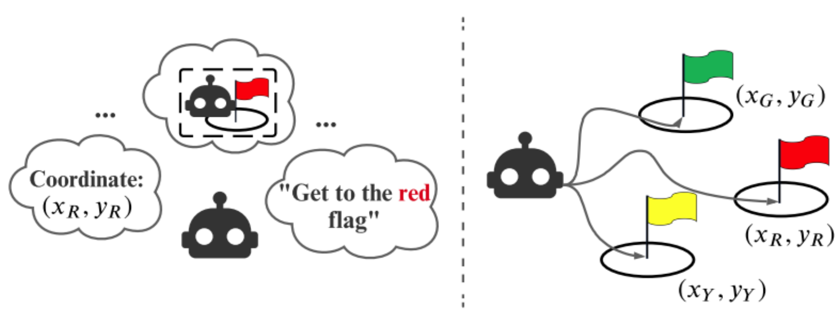

Figure 1: A typical picture of goals in GCRL from Liu *et al*. [23]. Commonly used goal representations in GCRL are vectors (embeddings), images and languages.

In this essay, we summarize various forms of goal representations. Basic representations usually use single perception, including vision, language, auditory sense, *etc*. Different from the former representations which can serve as inputs or outputs in end-to-end training, implicit embeddings and latent variables are also adopted to represent goals, mainly as intermediate variables. In the end, we will discuss the multi-modal representations, which is a combination of various perceptions.

2 Visual Perspective: Images as Goals

There is a famous theoretical proposal that "mental images are derived from goals" [8]. What we see in the image is determined by what our goal is; on the opposite perspective, an image description conveys enough information to represent a goal.

Much related work has focused on visual representations of goals. On the one hand, simulation environments including Atari games [4] and Minecraft [10, 15, 17] have become popular testbeds for image-based goal representations. On the other hand, image-goal-conditioned modeling are also brought to real world. For instance, Nair *et al*. [27, 28] first simulated a 7-dof Sawyer arm to reach goal positions, and then applied the agent to real world robotics control tasks with camera images.

Due to the high-dimensionality of images and visual inputs, many fantastic methods have been developed to tackle these problems in image-conditioned goals. For instance, variational auto-encoder (VAE) [19] is chosen to encode image representations into embeddings [18]. These embeddings serve as latent variables and participate in the subsequent calculations and modeling, such as representations of states in a GCRL setting [21].

3 Language Perspective: Texts as Goals

Natural languages are also widely used to represent goals. In comparison to images, texts have lower dimensionality and can convey more precise information using less spaces. However, text descriptions may not be as intuitive as images.

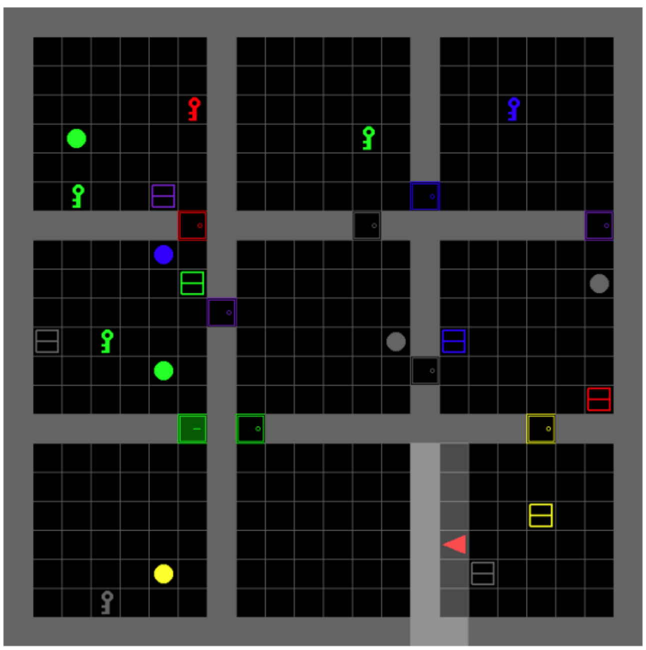

Figure 2: Boss level of BabyAI environment. The goal is to pick up the proper key and open the right door, all of which are described using natural language.

Natural languages are widely used in decision-making processes, such as Reinforcement Learning [25]. In most cases, the goal is an instruction sentence containing explicit verbs and objects, for instance, "go to blue torch" [5]. BabyAI [7] is a text-guided environment to train agents in a maze (Figure 2). SPiRL [30] trained robots with language instructions in a kitchen environment [14] (Figure 3).

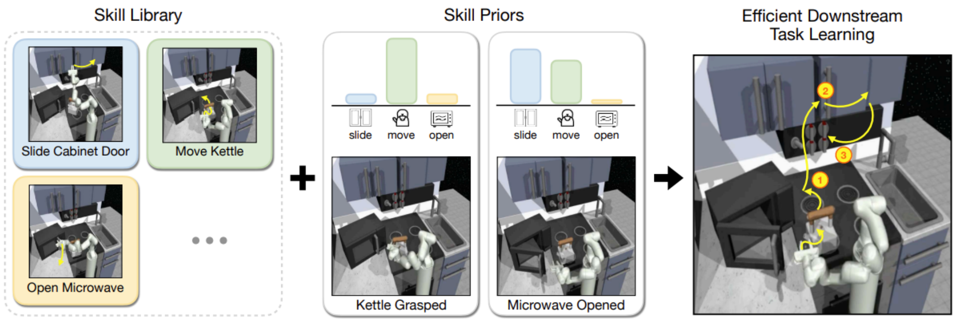

Figure 3: Intelligent agents learning to solve diverse tasks in a kitchen-like environment.

Other than reinforcement learning, another mainstream methods to solve language-guided goals is to use decision-transformer (DT) [6]. Many DT-based models [12, 20, 35] are skilled in tackling sequential texts, and are talented in making decisions accordingly.

4 Auditory Perspective: Sound as Goals

In comparison, fewer researches use audiroty sound as goals. Liu *et al*. [24] presented investigation on the mechanism between prediction versus goals in the context of adult Mandarin speakers' acquisition of non-native sounds, using an auditory feedback masking paradigm. Wang *et al*. [34] modeled auditory scene with analysis from a computational perspective. Fritz *et al*. [11] studied attention mechanism on sounds, in order to capture the auditory goals better. Stoilova *et al*. [32] used sound-cued reward to track goal-directed movement, shaping the behavioral adaptation.

5 Goal Representation in Latent Space

Latent goal representation is an aspect that can be applied to various domains due to its high implicity. Amado *et al*. [3] systematically introduced learning a goal recognition in latent space, and used LSTM-based method to encode these representations [2].

LatRec [1] managed to construct a method that combines goal recognition techniques from automated planning and deep auto-encoders, so as to carry out unsupervised learning to generate domain theories from data streams and use the resulting domain theories to deal with incomplete and noisy observations. LatRec has tackled the problem that a strong assumption has been made, in which there is a domain expert capable of building complete and correct domain knowledge to successfully recognize the goal of the agent. On another track, Hung *et al*. [16] solved path planning in robot manipulation through joint statistics. In this scene, goals are modeled as a distribution, according to which expected return are maximized.

Latent goal representation is commonly used in Goal-Conditioned Reinforcement Learning (GCRL). Nair *et al*.nair2018visual studied in a setting of visual reinforcement learning, but using imagined images as goals. They also use goal relabeling to impove sample efficiency. HIQL [29] proposed Hierarchical Implicit Q-Learning, a simple hierarchical method for offline goal-conditioned RL. Experiments were conducted on six types of state-based and pixel-based offline goal-conditioned RL benchmarks, and it was demonstrated that HIQL significantly outperformed previous offline goal-conditioned RL methods including GCBC [13], HGCBC [14] *etc*.

6 Multi-Modal Representations of Goals

Multi-modal learning is definitely a promising and rising field. However, the research in this area is still very preliminary, mainly focusing on combining vision and languages. Even so, this combination can bring much more extra information for agents in state-observation and decision-making.

Inspired by the progress in multi-modal learning, recent work has been exploring the possibility of representing goals in multi-modal forms. A typical version of multi-modal representations is to use images with text-labels or prompts. TransFuser *et al*. [31] was proposed as a novel Multi-Modal Fusion Transformer, to integrate image and LiDAR representations using attention. TransFuser was experimentally validated in urban settings involving complex scenarios using the CARLA urban driving simulator [9], and reached state-of-the-art (SOTA) performance. Luo *et al*. [26] proposed a novel instance-aware representation for lane representation by integrating the lane features and trajectory features. Then, a goal-oriented lane attention module is proposed to predict the future locations of the vehicle. It was shown that the proposed lane representation, together with the lane attention module, can be integrated into the widely used encoder-decoder framework to generate diverse predictions.

7 Conclusion

In this essay, we reviewed and summarized several different methods of goal representations. Basic ideas include represent goals with visual images, linguistic texts, and auditory sound. Goals represented with images can convey the most information, though lacking simplicity. Texts have excellent sequential properties and are simple enough, but cannot convey information as rich as images. Auditory ones are the least common, yet they are still important in tasks such as locating and positioning. In a more common setting - Goal-Conditioned RL, goals are more likely to be represented in latent space. These implicit embeddings are learned within the agent systems. Lastly, we mentioned the multi-modal representations of goals, though multi-modal learning is a preliminary field. However, multi-modal representations combine strengths and advantages of both vision and language, which can be of inspirations for future work.

References

[1] Leonardo Amado and Felipe Meneguzzi. Latrec: Recognizing goals in latent space (student abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 13747–13748, 2020. 3

[2] Leonardo Amado, Joao Paulo Aires, Ramon Fraga Pereira, Maurício C Magnaguagno, Roger Granada, and Felipe Meneguzzi. Lstm-based goal recognition in latent space. *arXiv preprint arXiv:1808.05249*, 2018. 3

[3] Leonardo Amado, Ramon Fraga Pereira, Joao Aires, Mauricio Magnaguagno, Roger Granada, and Felipe Meneguzzi. Goal recognition in latent space. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018. 3

[4] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym. *arXiv preprint arXiv:1606.01540*, 2016. 2

[5] Harris Chan, Yuhuai Wu, Jamie Kiros, Sanja Fidler, and Jimmy Ba. Actrce: Augmenting experience via teacher's advice for multi-goal reinforcement learning. *arXiv preprint arXiv:1902.04546*, 2019. 2

[6] Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Misha Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. Decision transformer: Reinforcement learning via sequence modeling. *Advances in neural information processing systems*, 34:15084–15097, 2021. 3

[7] Maxime Chevalier-Boisvert, Dzmitry Bahdanau, Salem Lahlou, Lucas Willems, Chitwan Saharia, Thien Huu Nguyen, and Yoshua Bengio. Babyai: A platform to study the sample efficiency of grounded language learning. *arXiv preprint arXiv:1810.08272*, 2018. 2

[8] Martin Conway, Kevin Meares, and Sally Standart. Images and goals. *Memory*, 12(4):525–531, 2004. 2

[9] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. Carla: An open urban driving simulator. In *Conference on robot learning*, pages 1–16. PMLR, 2017. 4

[10] Linxi Fan, Guanzhi Wang, Yunfan Jiang, Ajay Mandlekar, Yuncong Yang, Haoyi Zhu, Andrew Tang, De-An Huang, Yuke Zhu, and Anima Anandkumar. Minedojo: Building open-ended embodied agents with internet-scale knowledge. *Advances in Neural Information Processing Systems*, 35:18343–18362, 2022. 2

[11] Jonathan B Fritz, Mounya Elhilali, Stephen V David, and Shihab A Shamma. Auditory attention—focusing the searchlight on sound. *Current opinion in neurobiology*, 17(4):437–455, 2007. 3

[12] Hiroki Furuta, Yutaka Matsuo, and Shixiang Shane Gu. Generalized decision transformer for offline hindsight information matching. *arXiv preprint arXiv:2111.10364*, 2021. 3

[13] Dibya Ghosh, Abhishek Gupta, Ashwin Reddy, Justin Fu, Coline Devin, Benjamin Eysenbach, and Sergey Levine. Learning to reach goals via iterated supervised learning. *arXiv preprint arXiv:1912.06088*, 2019. 3

[14] Abhishek Gupta, Vikash Kumar, Corey Lynch, Sergey Levine, and Karol Hausman. Relay policy learning: Solving long-horizon tasks via imitation and reinforcement learning. *arXiv preprint arXiv:1910.11956*, 2019. 2, 3

[15] William H Guss, Brandon Houghton, Nicholay Topin, Phillip Wang, Cayden Codel, Manuela Veloso, and Ruslan Salakhutdinov. Minerl: A large-scale dataset of minecraft demonstrations. *arXiv preprint arXiv:1907.13440*, 2019. 2

[16] Chia-Man Hung, Shaohong Zhong, Walter Goodwin, Oiwi Parker Jones, Martin Engelcke, Ioannis Havoutis, and Ingmar Posner. Reaching through latent space: From joint statistics to path planning in manipulation. *IEEE Robotics and Automation Letters*, 7(2):5334–5341, 2022. 3

[17] Matthew Johnson, Katja Hofmann, Tim Hutton, and David Bignell. The malmo platform for artificial intelligence experimentation. In *Ijcai*, pages 4246–4247, 2016. 2

[18] Alexander Khazatsky, Ashvin Nair, Daniel Jing, and Sergey Levine. What can i do here? learning new skills by imagining visual affordances. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 14291–14297. IEEE, 2021. 2

[19] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 2

[20] Kuang-Huei Lee, Ofir Nachum, Mengjiao Sherry Yang, Lisa Lee, Daniel Freeman, Sergio Guadarrama, Ian Fischer, Winnie Xu, Eric Jang, Henryk Michalewski, et al. Multi-game decision transformers. *Advances in Neural Information Processing Systems*, 35:27921–27936, 2022. 3

[21] Lisa Lee, Ben Eysenbach, Russ R Salakhutdinov, Shixiang Shane Gu, and Chelsea Finn. Weakly-supervised reinforcement learning for controllable behavior. *Advances in Neural Information Processing Systems*, 33:2661–2673, 2020. 2

[22] Maria Legerstee, Joanne Barna, and Carolyn DiAdamo. Precursors to the development of intention at 6 months: understanding people and their actions. *Developmental psychology*, 36 (5):627, 2000. 1

[23] Minghuan Liu, Menghui Zhu, and Weinan Zhang. Goal-conditioned reinforcement learning: Problems and solutions. *arXiv preprint arXiv:2201.08299*, 2022. 1

[24] Xiaoluan Liu and Xing Tian. The functional relations among motor-based prediction, sensory goals and feedback in learning non-native speech sounds: Evidence from adult mandarin chinese speakers with an auditory feedback masking paradigm. *Scientific reports*, 8(1):11910, 2018. 3

[25] Jelena Luketina, Nantas Nardelli, Gregory Farquhar, Jakob Foerster, Jacob Andreas, Edward Grefenstette, Shimon Whiteson, and Tim Rocktäschel. A survey of reinforcement learning informed by natural language. *arXiv preprint arXiv:1906.03926*, 2019. 2

[26] Chenxu Luo, Lin Sun, Dariush Dabiri, and Alan Yuille. Probabilistic multi-modal trajectory prediction with lane attention for autonomous vehicles. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2370–2376. IEEE, 2020. 4

[27] Ashvin V Nair, Vitchyr Pong, Murtaza Dalal, Shikhar Bahl, Steven Lin, and Sergey Levine. Visual reinforcement learning with imagined goals. *Advances in neural information processing systems*, 31, 2018. 2

[28] Suraj Nair, Silvio Savarese, and Chelsea Finn. Goal-aware prediction: Learning to model what matters. In *International Conference on Machine Learning*, pages 7207–7219. PMLR, 2020. 2

[29] Seohong Park, Dibya Ghosh, Benjamin Eysenbach, and Sergey Levine. Hiql: Offline goal-conditioned rl with latent states as actions. *arXiv preprint arXiv:2307.11949*, 2023. 3

[30] Karl Pertsch, Youngwoon Lee, and Joseph Lim. Accelerating reinforcement learning with learned skill priors. In *Conference on robot learning*, pages 188–204. PMLR, 2021. 2

[31] Aditya Prakash, Kashyap Chitta, and Andreas Geiger. Multi-modal fusion transformer for end-to-end autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7077–7087, 2021. 3

[32] Vanya V Stoilova, Beate Knauer, Stephanie Berg, Evelyn Rieber, Frank Jäkel, and Maik C Stüttgen. Auditory cortex reflects goal-directed movement but is not necessary for behavioral adaptation in sound-cued reward tracking. *Journal of Neurophysiology*, 124(4):1056–1071, 2020. 3

[33] Athena Vouloumanos, Kristine H Onishi, and Amanda Pogue. Twelve-month-old infants recognize that speech can communicate unobservable intentions. *Proceedings of the National Academy of Sciences*, 109(32):12933–12937, 2012. 1

[34] DeLiang Wang. On ideal binary mask as the computational goal of auditory scene analysis. In *Speech separation by humans and machines*, pages 181–197. Springer, 2005. 3

[35] Mengdi Xu, Yikang Shen, Shun Zhang, Yuchen Lu, Ding Zhao, Joshua Tenenbaum, and Chuang Gan. Prompting decision transformer for few-shot policy generalization. In *international conference on machine learning*, pages 24631–24645. PMLR, 2022. 3