Revamp of Natural Language Processing using Reinforcement Learning

Mr. Kandi Gururaja Rao[1], Mrs. H. M. Shamitha[2], Dr. Yerriswamy T.[3], Mr. Venumadhava M.[4], Dr. Vijaya Kumar A. V.[5]

[1]Computer Vision Engineer, IT - Research & Development, Bizbox Publishing Pvt. Ltd, New Delhi, India, email: kgururajarao94@gmail.com

[2] Associate Professor, Department of Electronics & Communication, Proudadevaraya Institute of Technology, Karnataka, India, email: shamitha.hm@gmail.com

[3]Professor, Computer Science & Engineering, KLE Institute of Technology,Karnataka, India,

email: swamy1976ty@gmail.com

[4]Assistant Professor, Department of Electronics & Communication, Proudadevaraya Institute of Technology, Karnataka, India, email: venu.dsp@gmail.com

[5]Associate Professor, Department of Computer Science & Engineering, Proudadevaraya Institute of Technology, Karnataka, India, email: pdit.vijay@gmail.com

Abstract

Natural language processing (NLP) is a theory-motivated range of computational techniques for the automatic analysis and representation of human language. NLP research has evolved from the era of punch cards and batch processing (in which the analysis of a sentence could take up to 7 minutes) to the era of Google and the likes of it (in which millions of webpages can be processed in less than a second). NLP is one of the dominant fields in data mining. With the increasing importance of Big Data Analytics today, NLP plays a major role in acquiring relevant information of importance to business and intelligence. Millions of items are uploaded on Web every day, with relevant as well as irrelevant data. Information retrieval and extraction from reviews, comments, social media etc. by customers is a complex task since most of the information is in semi structured and unstructured form. Ambiguity of large corpora on Web underlines the need for decent and efficient data

There are no sources in the current document.a mining technique. Due to the recent progress in Deep Neural Networks, Reinforcement Learning (RL) has become one of the most important and useful technology. It is a learning method where a software agent interacts with an unknown environment, selects actions, and progressively discovers the environment dynamics. RL has been effectively applied in many important areas of real life. Reinforcement learning as an extension to neural networking which is widely used in gaming for the purpose of Natural Language Processing. We can utilize the reward-driven algorithm for better results. Many Natural Language Processing (NLP) tasks (including generation, language grounding, reasoning, information extraction, coreference resolution, and dialog) can be formulated as deep reinforcement learning (DRL) problems which are explained as the different models in this paper.

Keywords— *Natural Language Processing, Neural Networks, Deep Learning, Reinforcement Learning*

DOI Number: 10.14704/nq.2022.20.13.NQ88170 **Neuro Quantology 2022; 20(13):1360-1366**

I. INTRODUCTION

Natural Language Processing (NLP) [1]is the computerized approach to analyzing text that is based on both a set of theories and a set of technologies. And, being a very active area of research and development, there is not a single agreed-upon definition that would satisfy everyone, but there are some aspects, which would be part of any knowledgeable person's definition. [2] 'Human-like language processing' reveals that NLP is considered a discipline within Artificial Intelligence (AI). And while the full lineage of NLP does depend on several other disciplines, since NLP strives for human-like performance, it is appropriate to consider it an AI discipline. The goal of NLP as stated above is "to accomplish human-like language processing". The choice of the word 'processing' is very deliberate and should not be replaced with 'understanding'. For although the field of NLP was originally referred to as Natural Language Understanding (NLU) in the early days of AI, it is well agreed today that while the goal of NLP is true NLU, that goal has not yet been accomplished. Representation Learning, Reasoning and Scalability are three closely related research subjects in Natural Language Processing. [3] NLP enables computers to perform a wide range of natural language related tasks at all levels, ranging from parsing and part of-speech (POS) tagging, to machine translation and dialogue systems. Deep learning architectures and algorithms have already made impressive advances in fields such as computer vision and pattern recognition. Following this trend, recent NLP research is now increasingly focusing on the use of new deep learning methods. Machine learning and Statistical techniques are powerful analysis tools yet to be incorporated in the new multidisciplinary field diversely termed as natural language processing (NLP) or computational linguistic. The linguistic knowledge may be ambiguous or contains ambiguity; therefore, various NLP tasks are carried out in order to resolve the ambiguity in speech and language processing. The current prevailing techniques for addressing various NLP tasks as a supervised learning are hidden Markov models (HMM), conditional random field (CRF), maximum entropy models (MaxEnt), support vector machines (SVM), Naive Bays, and deep learning (DL).The goal of this paper is to provide brief overview of basic categories of linguistic knowledge, to discuss different existing machine learning models and their classification into different categories and finally to provide a comprehensive review of different state of the art machine learning models with the goal that new researchers look into these techniques and depending on these.

Deep Reinforcement Learning (DRL) is an emerging research area that involves intelligent agents that learn to reason in Markov Decision Processes (MDP). Recently, DRL has achieved many stunning breakthroughs in Atari games and the Go game. In addition, DRL methods have gained significantly more attentions in NLP in recent years, because many NLP tasks can be formulated as DRL problems that involve incremental decision making. DRL methods could easily combine embedding based representation learning with reasoning and optimize for a variety of non-differentiable rewards. There are various applications of RL, including games, in particular, AlphaGo, robotics, natural language processing, including dialogue systems, machine translation, and text generation, computer vision, neural architecture design, business management, finance, healthcare, Industry 4.0, smart grid, intelligent transportation systems, and computer systems[7]

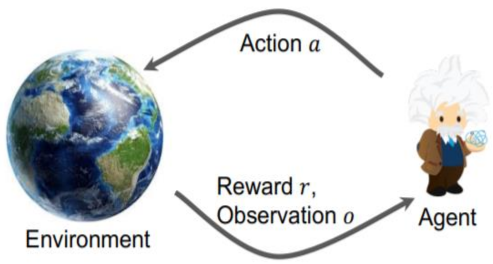

Figure 1: How Reinforcement Learning works

In recent years some researchers have explored the use of reinforcement learning (RL) algorithms as key components in the solution of various natural language processing tasks. For instance, some of these algorithms leveraging deep neural learning have found their way into conversational systems. This paper reviews the state of the art of RL methods for their possible use for different problems of natural language processing, focusing primarily on conversational systems, mainly due to their growing relevance. We provide detailed descriptions of the problems as well as discussions of why RL is well-suited to solve them. Also, we analyze the advantages and limitations of these methods. Finally, we elaborate on promising research directions in natural language processing that might benefit from reinforcement learning.

II. LITERATURE SURVEY

In this literature survey, we study the motivations, challenges, and recently developed algorithms of domain adaptation in the field of natural language processing (NLP). We first categorize recent work into three classes according to their methodologies, and then review and compare some preventative algorithms in each category. Finally, we summarize popular NLP applications involve domain adaptation.

A. Applications of Reinforcement Learning in Real World:

In this paper, Wahab Khan etal [6] has explained the Convolution Neural Network (CNN) and Recurrent Neural Network (RNN) are becoming more important for businesses due to their applications in Computer Vision (CV) and Natural Language Processing (NLP), Reinforcement Learning (RL) as a framework for computational neuroscience to model decision making process seems to be undervalued. Besides, there seems to be very little resources detailing how RL is applied in different industries. Despite the criticisms about RL's weaknesses, RL should never be neglected in the space of corporate research given its huge potentials in assisting decision making.

Therefore, this paper aims to

- Investigate the breadth and depth of RL applications in real world;
- View RL from different aspects; and
- Persuade the decision makers and researchers to put more efforts on RL research.

Recent neural models of dialogue generation offer great promise for generating responses for conversational agents, but tend to be shortsighted, predicting utterances one at a time while ignoring their influence on future outcomes. Modeling the future direction of a dialogue is crucial to generating coherent, interesting dialogues, a need which led traditional NLP models of dialogue to draw on reinforcement learning. In this paper, we show how to integrate these goals, applying deep reinforcement learning to model future reward in chatbot dialogue. The model simulates dialogues between two virtual agents, using policy gradient methods to reward sequences that display three useful conversational properties: informativity (non-repetitive turns), coherence, and ease of answering (related to forward-looking function). Evaluate model on diversity, length as well as with human judges, showing in figure 2 that the proposed algorithm generates more interactive responses and manages to foster a more sustained conversation in dialogue simulation. This work marks a first step towards learning a neural conversational model based on the long-term success of dialogues.

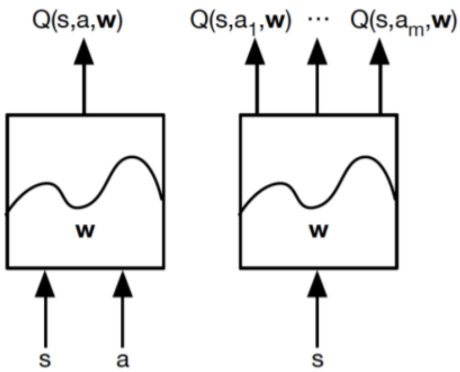

Figure 2: Forward-looking function of RL

B. Text Style Transfer

Text style transfer rephrases a text from a source style (e.g., informal) to a target style (e.g., formal) while keeping its original meaning. Despite the success existing works have achieved using a parallel corpus for the two styles, transferring text style has proven significantly more challenging when there is no parallel training corpus. In this paper, we address this challenge by using a reinforcement-learning-based generator-evaluator architecture. Our generator employs an attention-based encoder-decoder to transfer a sentence from the source style to the target style. Our evaluator is an adversarial trained style discriminator with semantic and syntactic constraints that score the generated sentence for style, meaning preservation, and fluency. Experimental results on two different style transfer tasks (sentiment transfer and formality transfer) show that our model outperforms state-of-the-art approaches. Furthermore, we perform a manual evaluation that demonstrates the effectiveness of the proposed method using subjective metrics of generated text quality is explained in the papar of Michael L Littman etal.[8]

State of the art NLP approaches generally adopt machine learning algorithms or more generally based on statistical machine learning. The prior attempts for linguistic processing were based on rules, rules are synthesized to carry out different NLP tasks. Unlike rule-based approach, ML approach automatically induces rules from training data. Machine learning algorithms usually consisted of intelligent modules which have capability to learn from historical data. Before ML approaches, NLP tasks are commonly carried out using rule based approaches. In Rule based approaches rules were constructed manually by linguistic experts or grammarians for particular tasks. Now a days various ML algorithms are in use for carrying various NLP tasks. The parameters of these algorithms is historical data from which features are synthesized, and this features based data latterly used in prediction or classification. Progressively, however, the Avant grade research in computational linguistic domain is based on machine learning models. In such a model, a real valued weight is attached to each input feature, which generate soft, probabilistic decisions. The benefit of such models is that: these models have the capability to represent a relation quality in different dimensions. Phases of RL is shown in Figure 3.

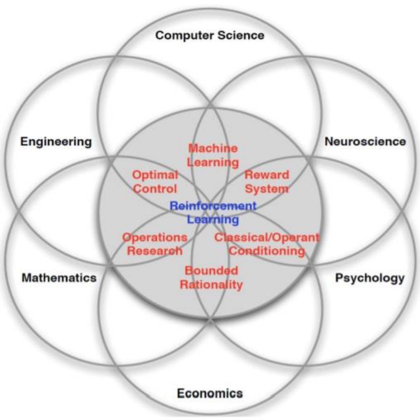

Figure 3: Phases of Reinforcement Learning (RL)

The state is a summary of experience

- At each step t (S_t) the agent:

Executes action a_t
Receives observation o_t
Receives scalar reward r_t

- The environment:

Receives action a_t
Emits observation a_{t-1}

$$S_t = f(o_1, r_1,\ a_1, \ldots, a_{t-1}, o_t, r_t) \text{--------(1)}$$

Represent policy by deep neural network that

max / ∅ - Maximum wave phase
a – observation
p – momentum,
s – step
r(a) – from the reward r and action a

$$\frac{max}{\emptyset}\ E_{a\sim p(a|\emptyset,s)}\ [r(a)|\emptyset, s] \text{-------(2)}$$

Experience Replay: remove correlations, build dataset from agent's own experience

A value function is a prediction of future reward: “How much reward will I get from action a in state s”

s – cycle or steps of the runs
a – action of run
r – reward from run

s1, a1, r2, s2
s2, a2, r3, s3
s3, a3, r4, s4
…..
St, at, rt+1, st+1

Table 1: Remove the correlations to build the dataset by the experience of the owner.

The state is a summary of experience of the NLP approach is given by the equation (1) and the same is represented by the seep neural network by the equation (2). By removing the correlations to build the dataset from the agent’s experience is given in the above tabular column.

III. METHODOLOGY

In reviewing efforts that integrate language in RL we highlight work that develops tools, approaches, or insights that we believe may be particularly valuable for improving the generalization or sample efficiency of learning agents using natural language. As illustrated in

Neuro Quantology | October 2022 | Volume 20 | Issue 13 | Page 1360-1366 | doi: 10.14704/nq.2022.20.13.NQ88170

Figure 4, we separate the literature into **language-conditional RL** (in which interaction with language is necessitated by the problem formulation itself) and **language-assisted RL** (in which language is used to facilitate learning). The two categories are not mutually exclusive, in that for some language conditional RL tasks, NLP methods or additional textual corpora are used to assist learning

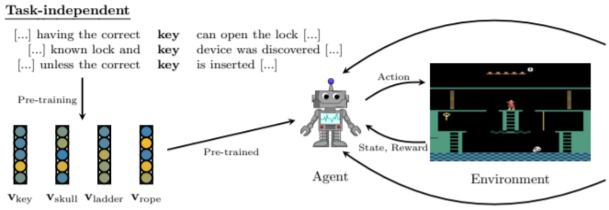

Figure 4: Illustration of different roles and types of natural language information in reinforcement learning

- **Language-conditional RL**

We first review literature for tasks in which integrating natural language is unavoidable, i.e., when the task itself is to interpret and execute instructions given in natural language, or natural language is part of the state and action space. Which approaches to such tasks which can also be improved by developing methods that enable transfer from general and task-specific textual corpora. Methods developed for language conditional tasks are relevant for language assisted RL as they both deal with the problem of grounding natural language sentences in the context of RL.

- **Language-assisted RL**

In this section, we consider work that explores how knowledge about the structure of the world can be transferred from natural language corpora and methods into RL tasks, in cases where language itself is not essential to the task. Textual information can assist learning by specifying informative features, annotating states or entities in the environment, or describing subtasks in a multitask setting. In most cases covered here, the textual information is task specific.

RL, known as a semi-supervised learning model in machine learning, is a technique to allow an agent to take actions and interact with an environment so as to maximize the total rewards. RL is usually modeled as a Markov Decision Process (MDP).

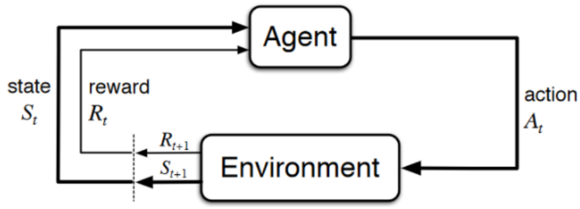

Figure 5: The RL agent-environment interaction

To advance the field, we argue that more research effort should be spent on learning from naturally occurring text corpora in contrast to instruction following. While learning from unstructured and descriptive text is particularly difficult, it has a much greater application range and potential for impact. Moreover, we argue for the development of more diverse environments with real-world semantics figure 5. The tasks used so far use small and synthetic language corpora and are too artificial to significantly benefit from transfer from real-world textual corpora. In addition, we emphasize the importance of developing standardized environments and evaluations for comparing and measuring progress of models that integrate natural language into RL agents.

We believe that there are several factors that make focusing such efforts worthwhile now:

(i) Recent progress in pre-training language models,
(ii) General advances in representation learning, as well as
(iii) Development of tools that make constructing rich and challenging RL environments easier. Some significant work, especially in language assisted RL, has been done prior to the surge of deep learning methods.

Some of the applications which uses NL and RL

Resources management in computer clusters- Designing algorithms to allocate limited resources to different tasks is challenging and requires human-generated heuristics. The paper “Resource Management with Deep Reinforcement Learning” [2] showed how to use RL to automatically learn to allocate and schedule computer resources to waiting jobs, with the objective to minimize the average job slowdown.

Traffic Light Control-In the paper “Reinforcement learning-based multi-agent system for network traffic signal control”, researchers tried to design a traffic light controller to solve the congestion problem.

Tested only on simulated environment though, their methods showed superior results than traditional methods and shed a light on the potential uses of multi-agent RL in designing traffic system.

Robotics-There are tremendous work on applying RL in Robotics. Readers are referred to [10] for a survey of RL in Robotics. In particular, [11] trained a robot to learn policies to map raw video images to robot's actions. The RGB images were fed to a CNN and outputs were the motor torques. The RL component was the guided policy search to generate training data that came from its own state distribution.

Web System Configuration-There are more than 100 configurable parameters in a web system and the process of tuning the parameters requires a skilled operator and numerous trail-and-error tests. The paper "A Reinforcement Learning Approach to Online Web System Auto-configuration" [5] showed the first attempt in the domain on how to do autonomic reconfiguration of parameters in multi-tier web systems in VM-based dynamic environments.

IV. CONCLUSION

The currently predominant way RL agents are trained restricts their use to environments where all information about the policy can be gathered from directly acting in and receiving reward from the environment. Utilizing natural language in RL agents could drastically change this by transferring knowledge from natural language corpora to RL tasks, as well as between tasks, consequently unlocking RL for more diverse and real-world tasks. While there is a growing body of papers that incorporate language into RL, most of the research effort has been focused on simple RL tasks and synthetic languages, with highly structured and instructive text. To realize the potential of language in RL, we advocate for more research into learning from unstructured or descriptive language corpora, with a greater use of NLP tools like pre-trained language models. Such research also requires development of more challenging environments that reflect the semantics and diversity of the real world. This paper also showed some of the examples of RL applications in various industries. They should not limit your RL use case and as always, you should use first principle to understand the nature of RL and your problem.

REFERENCES

[1] Sharma, A.R. and Kaushik, P., 2017, May. Literature survey of statistical, deep and reinforcement learning in natural language processing. In 2017 International conference on computing, communication and automation (ICCCA) (pp. 350-354). IEEE.

[2] Young, T., Hazarika, D., Poria, S. and Cambria, E., 2018. Recent trends in deep learning based natural language processing. ieee Computational intelligenCe magazine, 13(3), pp.55-75.

[3] Naeem, M., Rizvi, S.T.H. and Coronato, A., 2020. A gentle introduction to reinforcement learning and its application in different fields. IEEE Access.

[4] Madureira, B. and Schlangen, D., 2020. An overview of natural language state representation for reinforcement learning. arXiv preprint arXiv:2007.09774.

[5] Cambria, Erik; White, Bebo 2014. Jumping NLP Curves: A Review of Natural Language Processing Research [Review Article]. IEEE Computational Intelligence Magazine, 9(2), 48–57. doi:10.1109/mci.2014.23072

[6] A survey on the state-of-the-art machine learning models in the context of NLP Wahab Khan Ali Daud, Jamal A. Nasir1, Tehmina Amjad

[7] Deep Reinforcement Learning for Dialogue Generation Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, Dan Jurafsky

[8] Monica Babes, Vukosi Marivate, Kaushik Subramanian, and Michael L Littman. Apprenticeship learning about multiple intentions. In ICML, 2011.

[9] Reinforcement Learning: An Introduction. By Richard S. Sutton and Andrew G. Barto Second Edition, in progress MIT Press, Cambridge, MA, 2017

[10] Rishabh Agarwal, Chen Liang, Dale Schuurmans, and Mohammad Norouzi. Learning to generalize from sparse and

underspecified rewards. CoRR, abs/1902.07198, 2019.

[11] Maxime ChevalierBoisvert, 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. In ICLR, 2019.

[12] Michele Banko, Michael J Cafarella, Stephen Soderland, Matthew Broadhead, and Oren Etzioni. Open information extraction from the web. In IJCAI, 2007

[13] Karthik Narasimhan, Tejas D. Kulkarni, and Regina Barzilay. Language understanding for text-based games using deep reinforcement learning. In EMNLP, 2015

[14] Survey on reinforcement learning for language processing by Victor Uc-Cetina, Nicolas Navarro-Guerrero, Anabel Martin-Gonzalez, Cornelius Weber, Stefan Wermter