Recognition Method of Important Words in Korean Text Based on Reinforcement Learning

Feiyang Yang, Yahui Zhao[✉], and Rongyi Cui

Department of Computer Science and Technology, Yanbian University,
977 Gongyuan Road, Yanji 133002, China
903873610@qq.com

Abstract. The manual labeling work for constructing the Korean corpus is too time-consuming and laborious. It is difficult for low-minority languages to integrate resources. As a result, the research progress of Korean language information processing is slow. From the perspective of representation learning, reinforcement learning was combined with traditional deep learning methods. Based on the Korean text classification effect as a benchmark, and studied how to extract important Korean words in sentences. A structured model Information Distilled of Korean (IDK) was proposed. The model recognizes the words in Korean sentences and retains important words and deletes non-important words. Thereby transforming the reconstruction of the sentence into a sequential decision problem. So you can introduce the Policy Gradient method in reinforcement learning to solve the conversion problem. The results show that the model can identify the important words in Korean instead of manual annotation for representation learning. Furthermore, compared with traditional text classification methods, the model also improves the effect of Korean text classification.

Keywords: Reinforcement learning · Attention mechanism · Korean natural language processing · Structure discovery

1 Introduction

The languages of ethnic minorities have created the diversity of Chinese characters and are an important part of Chinese characters, providing important support for the development of national culture. However, the research on Korean natural language processing in my country is still in the development stage, and the related research is still relatively lagging behind South Korea and North Korea [2]. For manual annotation of Korean sentences, the structure division requires a lot of energy and time. So for this problem, we associate the method of representation learning. Representation learning has been widely used in text classification, sentiment analysis, language reasoning and other fields in recent

years. It is a basic problem in the field of artificial intelligence, and it is particularly important in the field of natural language processing. Therefore, we use this method as the core logic of the model, aiming at Korean text, identifying important words and performing sentence classification tasks on newly constructed sentences. The resulting structural representation does not require manual annotation, greatly reducing manpower and scientific research resources.

In order to find important Korean words in sentences, we use the effect of text classification as feedback in reinforcement learning. The current mainstream text classification models are roughly divided into four types: bag-of-words model, sequence model, structure representation model, and attention model. The bag-of-words representation model often ignores the order of words, such as deep average networks, self-encoders [5]; the sequence representation model often only considers the words themselves, but ignores the phrase structure, such as CNN, RNN and other neural network models [14]; structural representation models often rely on pre-specified parse trees to construct structured representations, such as Tree-LSTM, recursive autoencoders [17]; representation models based on attention mechanisms need to use input words or sentences The attention scoring function is used to construct a representation, such as Self-Attention [15], and the effect is very dependent on the reliability of scoring. In the existing structured representation model, the structure can be provided as input, or it can be predicted using the supervised method of explicit tree annotations, but few studies have studied the representation with automatically optimized structure. *Yogatama et al.* proposed to construct a binary tree structure for sentence representation only under the supervision of downstream tasks, but this structure is very complicated and the depth is too large, resulting in unsatisfactory classification performance [16]. *Chung et al.* proposed a hierarchical representation model to capture the latent structure of sequences with latent semantics, but the structure can only be found in the hidden space [3]. *Tianyang et al.* proposed a method that combines the strategy gradient method in reinforcement learning with the LSTM model in deep learning. The effect of text classification is used as the baseline for reinforcement learning to carry out unsupervised structuring, and its structuring effect is closer to Human, and the classification effect is significantly better than other mainstream models [12].

Inspired by *Tianyang et al.*, we propose a method that incorporates reinforcement learning. By identifying the structure related to the task, it does not require explicit structural annotations to construct a sentence representation. Among them, the structure discovery problem is transformed into a sequence decision problem. Using the policy gradient method in reinforcement learning (Policy Gradient), the value of the delayed reward function is used to guide the self-discovery of the structure. The definition of the reward function is expressed in the same text according to the structure. The classification effect of the classifier is derived, Each time the structured representation obtained needs to be used after all sequential decisions have been made, The model incorporates an attention mechanism and a baseline of reinforcement learning convergence on the basis of predecessors to optimize it. The main purpose of the model IDK we

designed is to delete the unimportant words in the sentence, retain the words most relevant to the task, and construct a new sentence representation, in which the strategy network, the structured representation, and the classification network are seamlessly integrated. The strategy network defines the strategy used to discover the structure. The classification network calculates the classification accuracy based on the structured sentence representation, and passes the value of the reward function to the strategy network to promote the self-optimization of the entire network model.

2 Strategy Network Based on Reinforcement Learning Combined with Attention Mechanism

The core idea of the strategy network is the Policy Gradient method, which is different from the traditional method. Instead of backpropagating through errors, the observation reward value is used to enhance or weaken the possibility of selecting actions. That is, the probability that a good action will be selected next time will increase; the probability that a bad action will be selected next time will decrease. A complete strategy represents a sequence of actions taken in each state in a round, The cumulative sum of the revenue generated by each action represents the round reward value. We use a random strategy $\pi\left(a_t|\mathbf{s_t};\Theta\right)$. And use the reward generated by each round delay to guide strategy learning, For each state of the structured representation model generated each time, different actions are sampled. First, all the words of the entire sentence must be sampled for action, so as to determine the actions of all states corresponding to a sentence, Secondly, the determined action sequence is passed into the representation model to generate a new structured representation; then the generated structured representation is passed into the classification network, and this representation model is used to calculate the classification accuracy $P(y|X)$, Finally, the calculated reward is used for strategy learning. In the loop iteration, a better strategy is found, so that a better structured representation is obtained, and then a better classification effect is obtained. The strategy is defined as follows:

$$\pi\left(a_t|\mathbf{s_t};\Theta\right) = \sigma\left(\mathbf{W}*\mathbf{s}_t + \mathbf{b}\right) \tag{1}$$

Which a_t represents the probability of choosing at; σ represents the sigmoid function; Θ represents the parameters of the strategy network. During the training, actions are sampled according to the probability in Eq. 1. During the test, the action with the highest probability will be selected to achieve a better classification effect.

$$a_t^* = \mathrm{argmax}_a \pi(a\,|\mathrm{s_t};\Theta\,) \tag{2}$$

When all actions are sampled by the strategy network, the structured representation of the sentence is determined by the representation model, and the determined representation model is passed to the classification network to obtain, where y is the classification label, reward will be calculated from the predicted

distribution, and There are also factors for thinking about the trend of structural choices. Therefore, the Policy Gradient method in the reinforcement learning algorithm is used to optimize the parameters of the strategy network [6], so as to maximize the expected return, as shown in Eq. 3.

$$
\begin{aligned}
J(\Theta) &= \sum_{\mathbf{s_1}a_1\cdots\mathbf{s_L}a_L} P_\Theta\left(\mathbf{s_1}a_1\cdots\mathbf{s_L}a_L\right)R_L \\
&= \sum_{\mathbf{s_1}a_1\cdots\mathbf{s_L}a_L} p\left(\mathbf{s_1}\right)\prod_t \pi_\Theta\left(a_t|\mathbf{s_t}\right)p\left(\mathbf{s_{t+1}}|\mathbf{s_t},a_t\right)R_L \\
&= \sum_{\mathbf{s_1}a_1\cdots\mathbf{s_L}a_L}\prod_t \pi_\Theta\left(a_t|\mathbf{s_t}\right)R_L
\end{aligned}
\tag{3}
$$

The reward calculation is only for one round, because the state at step $t+1$ is completely determined by the state at step t, so the probability sum is 1. Through the likelihood ratio technique, the following gradient update strategy network [11] is finally used, where N represents the round number. In the iterative process of reinforcement learning, the variance is generally large. If the loss value is always positive, the direction of the iteration is easy to move toward. It has been proceeding in the wrong direction, so the introduction of b as a baseline can accelerate convergence, as shown in Eq. 4.

$$
\nabla_\Theta J(\Theta) = -\frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{L}(R_L - b)\nabla_\Theta \log \pi_\Theta\left(a_t|\mathbf{s_t}\right)
\tag{4}
$$

On this basis, the attention mechanism is introduced into the strategy network, and the Encoder-Decoder framework is adopted. Use Bi-LSTM [4] as the encoder model, LSTM as the decoder model, and the output of the structured representation model as the input of the strategy network, because the core logic of the attention model is from focusing on the whole to focusing on the core. The purpose of this article is the same, so the combination of reinforcement learning and Soft Attention mechanism to make up for the shortcomings of the predecessors in the traditional attention model in the text classification process, relying heavily on the scoring function [1]. After introducing the attention mechanism, the corresponding actions in each state are output as shown in Fig. 1.

$$
S_t = f(S_{t-1}, Y_{t-1}, C_t)
\tag{5}
$$

$$
C_t = \sum_{j=1}^{T}\alpha_{tj}h_j
\tag{6}
$$

$$
\alpha_{tj} = \frac{exp(e_{tj})}{\sum_{k=1}^{T}exp(e_{tk})}
\tag{7}
$$

$$
e_{tj} = g(S_{t-1}, h_j)
\tag{8}
$$

Where h_j is the hidden vector of the input, f is the tanh activation function; C is the attention distribution; and α_{tj} is the attention obtained by each input.
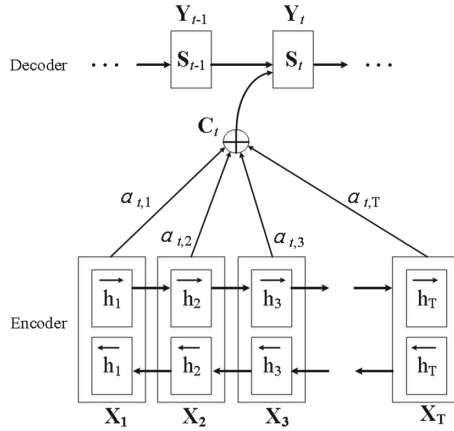
Fig. 1. Soft attention mechanism

After introducing the attention mechanism, the global observation can be better, so that the generated action sequence is optimized in two aspects of the strategy gradient and the attention mechanism, thereby improving the model effect.

3   Information Distilled of Korean (IDK)

3.1   The Main Idea of the Model

Our ultimate goal is to reconstruct more concise Korean sentences by finding important, task-related words, and at the same time get a structural representation for Korean text classification, and improve text classification through optimized structured representation. While the text classification has been improved, the structured representation has also been optimized, and the two promote each other. The model consists of three parts: strategy gradient network, structured representation model, classification network. The strategy network adopts a random strategy to sample the actions corresponding to each state, sampling until the end of the sentence, and generating a sequence of actions for the current sentence. Then the structured representation model converts the action sequence into a structured representation. Based on this idea, the IDK model is proposed. The classification network classifies based on the obtained structured representation and provides the reward function calculation for the strategy network. Since a complete structured representation can be given to calculate the reward of the current structured representation, this process can be solved by the Policy Gradient method. The specific model is shown in Fig. 2.

The model is interleaved by three parts. The state representation of the strategy network comes from the structured model. The structured model is generated by the action sequence of the strategy network and the input of the sentence. The classification network is classified and predicted by the resulting
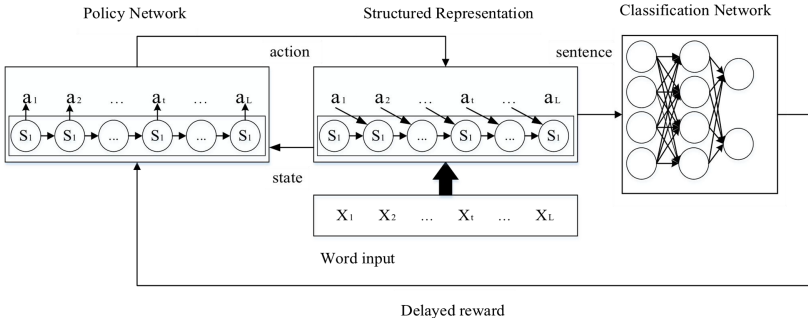
Fig. 2. Network model structure diagram based on reinforcement learning.

structured model. Strategy The network obtains the reward function value from the classification effect obtained by the classification network, thereby guiding the strategy to learn a better structured representation.

3.2 Model Specific Construction

Inspired by the ID-LSTM of *Tianyang et al.*, an attention mechanism was introduced into the original strategy network to double optimize the action sequence. The main idea of IDK proposed in our thoughts is to build a structured representation of a sentence by extracting important words and deleting irrelevant words in the sentence. In the well-known Chinese and English text processing tasks, there are many examples such as: "with", "and", "in", "of" and other stop words, such stop words rarely help complete text processing tasks, so it is necessary to refine important features in sentences. Different from the traditional method, this method does not create a stop word list, and deletes all stop words together. Because many stop words often constitute a special phrase structure, combing the logical relationship between the context and deleting it directly without filtering, it will cause the loss of language content and semantic information, so this method is chosen in our thoughts to purify the final representation form, thus Concentrate sentences to enhance the effectiveness of downstream classification tasks.

The IDK model converts the sequence of actions transferred from the strategy network into a structured representation of sentences, Given a sentence X shaped like $X = x_1 x_2 \cdots x_L$, After the sentence X is transferred to the strategy network, each action a_i corresponding to the word position x_i is selected from keeping the current word or deleting the current word, which satisfies the following rules:

$$\mathrm{S_t, C_t,} = \begin{cases} \mathrm{S_{t-1}, C_{t-1},} & a_t = Delete \\ \Phi\left(\mathrm{S_{t-1}, C_t, Y_{t-1}}\right), & a_t = Retain \end{cases} \tag{9}$$

The Φ represents the function of the entire model (including gating unit and update function), S is the hidden state corresponding to the Decoder cell unit; Y is the output corresponding to the hidden state of the cell unit; C is the

hidden state distribution of the Encoder cell unit; When deleting a word, the storage unit and hidden state attention distribution of the current position will be copied from the previous position.

For classification, the last hidden state of the IDK model is used as the input of the classification network, where $\mathbf{W_s} \in R^{d \times K}$, $\mathbf{b_s} \in R^K$ is the parameter of the classification network, d is the dimension of the hidden state, is the label of the category, K is the number of classification clusters, the classification network is based on the IDK model The obtained structured representation produces a probability distribution on the class label, as shown in Eq. 10.

$$P(y\,|X\,) = \mathrm{softmax}(\mathbf{W_s}S_L + b_s) \tag{10}$$

To calculate reward, take the logarithm of the output probability calculated by the classification network in Eq. 10, as shown in Eq. 11, where c_g stands for classification label. In order to make the model use as few words as possible, the two items in the formula are controlled by calculating the ratio of the number of deleted words in the sentence to the length of the sentence. Maintain accuracy and balance the two effects of using a few words, Where L' represents the number of deleted words, and γ represents the hyperparameter between 0 and 1 that balances the two terms.

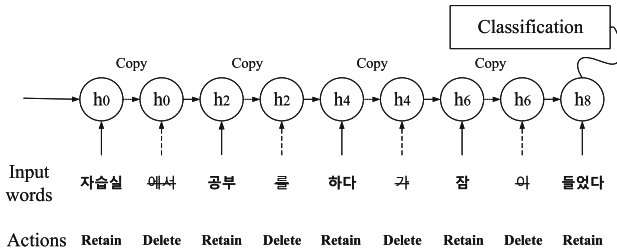$$R_L = \log P\,(c_g|X) + \gamma L'/L \tag{11}$$

Fig. 3. ID-Korean model

As shown in Fig. 3, after word recognition is performed on Korean text, only important words are retained, and the ratio of the amount retained to the number deleted is controlled by the reward function.

4   Experimental Results and Analysis

4.1   Data Set Description

The data set used in the experiments in this article comes from the corpus constructed by the laboratory to undertake the "China-Korea Science and Technology Information Processing Comprehensive Platform" project. It is further

organized into a corpus composed of abstracts of Korean scientific and techno-
logical literature. There are about 30,000 documents, divided into 13 categories
such as animals, oceans, and aerospace. Each Category randomly selects docu-
ments according to a 7:3 ratio to form a training set and a test set [9,13]. The
details of the data set are shown in Table 1.

Table 1. Data set introduction

Category	Number of entries	Category	Number of entries
Animal	4582	Botany	6172
Microorganism	5472	Biotechnology	1215
Biomedical science	2752	Climate	708
The marine environment	810	Geology	1735
Marine technology	819	Materials engineering	781
Measurement technology	1728	Aerospace	4436
Others	1478		

For Korean corpora, sentences are composed of phrases separated by spaces,
and these phrases are usually followed by auxiliary words or endings. According
to the grammatical characteristics of Korean, In the preprocessing process, the
Hannanum word segmentation system developed by the Korea University of
Science and Technology is used to cut out the auxiliary words and endings in
the phrase, and restore the predicate to the word itself.

4.2   Model Training

When training a classification network, a cross-entropy loss function is used,
in which the probability distribution of the ground truth of the corresponding
sentence is coded by one-hot, as shown in Eq. 12.

$$\mathcal{L} = - \sum_{X \in \mathcal{D}} \sum_{y=1}^{K} \hat{p}(y, X) \log P(y|X) \tag{12}$$

GloVe training is used to initialize the word vector in the representation
model [10], the dimension is set to 256 dim, and it is updated together with
other parameters. When using gradient descent to update parameters, the speed
of model learning depends on the learning rate and partial derivative value,
To smooth the update of Policy Gradient, multiply the suppression factor γ
by Eq. 4 and set it to 0.1, γ is set to 0.2 in the IDK, Eq. 11, In the training
process, the Adam optimizer [7] is used to optimize the parameters, the learning
rate is 0.0005, the Dropout tailoring is used before the classification network
classification, the probability is 0.5, and the mini-batch is 5.

In the model training process, the classification accuracy rate using the IDK model changes with the number of iterations as shown in Fig. 4. The text classification accuracy rate is about 68% at the beginning of the training, and the accuracy rate increases as the number of iterations increases. When the number of iterations is between 400 and 600, the classification accuracy of the model rises fastest. After 800 iterations, the classification accuracy of multilingual text tends to be stable, indicating that the training of the neural network model has converged. At this time, the text classification of the IDK model The accuracy rate reached 83.23%.
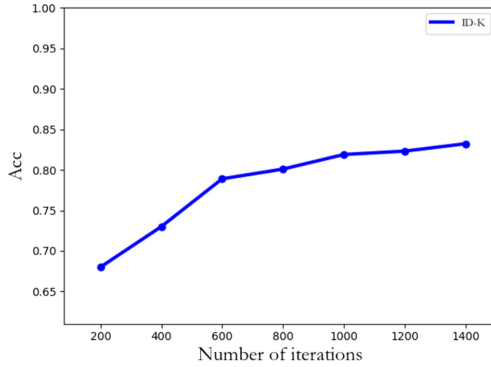

Fig. 4. Soft attention mechanism

4.3   Comparative Experiment

In the comparative experiment, a variety of baselines were selected: basic neural network model CNN without specific structure; LSTM; Bi-LSTM; attention model Self-Attention [8]. The dimension of the word vector used by these baselines is the same as this article, and the effect is shown in Table 2.

Table 2. Accuracy under different classifiers

Models	ACC	Models	ACC
CNN	78.5	T-BLSTM-CNN	81.68
LSTM	74.6	Self-Attention	82.91
Bi-LSTM	78.14	**IDK**	**83.23**

As shown in Table 2, the classification effect shows that: in different models, our method performs well in classification. When comparing with previous

methods, we combine reinforcement learning and attention model to use a self-discovery structure and Optimize the structured representation model for text classification, Different from the predecessors who only focused on the sequence model and its optimization, this paper designs the model from two aspects of reinforcement learning and attention. Its classification effect also proves the effectiveness and necessity of representation learning and text structuring.

4.4   Examples of Structured Presentation Results

Original	도시는 인류 활동의 영향을 가장 크게 받는 지구의 표면 이고 도시 시스템의 탄소 순환은 세계 및 지역의 탄소 순환에서 중요한 위치와 역할을 한다（Cities are the surface of the earth, which is most affected by human activity, and the carbon cycle of urban systems plays an important position and role in the carbon cycle of the world and regions.）
IDK	도시는 인류 활동의 영향을 가장 크게 받는 지구의 표면 이고 도시 시스템의 탄소 순환은 세계 및 지역의 탄소 순환에서 중요한 위치와 역할을 한다
Original	곤충 병원체에서 직접 샘플을 채취하고 검출 및 정량화하는 것은 곤충 유행병학 조사에서 병원체 풍도를 직접 반영할 수 있다（Taking, detecting, and quantifying samples directly from insect pathogens can directly reflect pathogen wind levels in insect epidemiological surveys.）
IDK	곤충 병원체에서 직접 샘플을 채취하고 검출 및 정량화하는 것은 곤충 유행병학 조사에서 병원체 풍도를 직접 반영할 수 있다

Fig. 5. Structural representation example

The specific structured example is shown in Fig. 5. In the IDK model, the strike through indicates the word to be deleted on the original text. The Korean text's mood words, auxiliary words, and some adjectives are deleted, and important part of the nouns are retained. The larger the model segmentation structure is, the closer it is to manual annotation, and the original text is more useful for downstream text classification tasks after being structured.

5   Conclusion

We combined reinforcement learning methods to learn Korean sentence representations by finding important words related to the task. In the framework of reinforcement learning and attention mechanism, this paper uses the IDK model, which is used to extract task-related words and express them in purified sentences. Among them, reinforcement learning uses the accuracy rate of text classification as a baseline to optimize the action sequence, and the action sequence can generate a text structure representation that is more suitable for

classification. An attention mechanism is introduced in the process of action sequence generation to compensate for the variance of the reinforcement learning method. The disadvantage of being too large and difficult to fit, compared with the traditional attention model, not only has its advantages of taking into account the overall situation, but also adds a more ingenious way to improve the accuracy of downstream tasks, and the experimental results have performed well. Experiments show that our method can find important words related to tasks without explicit structure annotation. The model not only improves the effect of Korean text classification, but also works well in the task of processing Korean text important word recognition.

References

1. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. In: International Conference on Machine Learning. arXiv:1409.0473v7 (2015)
2. Bi, Y.: A research on Korean natural language processing. J. Chin. Inf. Process. **25**(06), 166–169+182 (2011)
3. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. In: NIPS 2014 Workshop (2014)
4. Graves, A., Schmidhuber, J.: Frame wise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw. **18**(5–6), 602–610 (2005)
5. Joulin, A., Grave, E., Bojanowski, P., Mikolov, T.: Bag of tricks for efficient text classification. In: The European Chapter of the ACL (EACL), pp. 427–431 (2017)
6. Keneshloo, Y., Ramakrishnan, N., Reddy, C.K.: Deep transfer reinforcement learning for text summarization. Society for Industrial and Applied Mathematics. arxiv:1810.06667v2 (2019)
7. Kingma, D.P., Ba, L.J.: A method for stochastic optimization. In: International Conference on Learning Representations (2015)
8. Lin, Z., et al.: A structured self-attentive sentence embedding. In: International Conference on Learning Representations (2017)
9. Tian, M., Zhao, Y., Cui, R.: Identifying word translations in scientific literature based on labeled bilingual topic model and co-occurrence features. In: Sun, M., Liu, T., Wang, X., Liu, Z., Liu, Y. (eds.) CCL/NLP-NABD - 2018. LNCS (LNAI), vol. 11221, pp. 76–87. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01716-3_7
10. Pennington, J., Socher, R., Manning, C.D.: Glove: global vectors for word representation. In: EMNLP 2014, pp. 1532–1543 (2014)
11. Sutton, R.S., McAllester, D.A., Singh, S.P., Mansour, Y.: Policy gradient methods for reinforcement learning with function approximation. In: NIPS 2000, pp. 1057–1063 (2000)
12. Zhang, T., Huang, M., Zhao, L.: Learning structured representation for text classification via reinforcement learning. In: The Association for the Advance of Artificial Intelligence (2018)

13. Meng, X., Cui, R.: Multilingual text classification method based on bidirectional long-short memory unit and convolutional neural network. Comput. Appl. Res. **132**(04), 1–6 (2019)
14. Kim, Y.: Convolutional neural networks for sentence classification. In: The 2014 Conference on Empirical Methods in Natural Language Processing, pp. 1746–1751 (2014)
15. Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., Hovy, E.: Hierarchical attention networks for document classification. In: Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 1480–1489 (2016)
16. Yogatama, D., Blunsom, P., Dyer, C., Grefenstette, E., Ling, W.: Learning to compose words into sentences with reinforcement learning. In: International Conference on Learning Representations (2017)
17. Zhu, X., Sobihani, P., Guo, H.: Long short-term memory over recursive structures. In: International Conference on Machine Learning, pp. 1604–1612 (2015)