

SOLVING THE CHINESE PHYSICAL PROBLEM BASED ON DEEP LEARNING AND KNOWLEDGE GRAPH

Mingchen Li¹, Zili Zhou², Yanna Wang¹

¹College of Physics and Engineering, Qufu Normal University, Qufu 273165,
PR China

²School of Software Engineering, Qufu Normal University, Qufu Shandong
273165, PR China

ABSTRACT

In recent years, problem solving, automatic proof and human-like test-tasking have become a hot spot of research. This paper focus on the study of solving physical problem in Chinese. Based on the analysis of physical corpus, it is found that the physical problem are made up of n-tuples which contain concepts and relations between concepts, and the n-tuples can be expressed in the form of UP-graph (The graph of understanding problem), which is the semantic expression of physical problem. UP-graph is the base of problem solving which is generated by using physical knowledge graph (PKG). However, current knowledge graph is hard to be used in problem solving, because it cannot store methods for solving problem. So this paper presents a model of PKG which contains concepts and relations, in the model, concepts and relations are split into terms and unique IDs, and methods can be easily stored in the PKG as concepts. Based on the PKG, DKP-solving is proposed which is a novel approach for solving physical problem. The approach combines rules, statistical methods and knowledge reasoning effectively by integrating the deep learning and knowledge graph. The experimental results over the data set of real physical text indicate that DKP-solving is effective in physical problem solving.

KEYWORDS

Knowledge Graph, Deep Learning, Problem Solving, & Physical Problem

1. INTRODUCTION

Developing machine solver [1] has become an active research topic since the 1960s and recently it has developed rapidly under the influence of fast advance of artificial intelligent (AI), Human-Computer Interaction [2], machine learning [3, 4], and pattern recognition. It can be applied to various courses, for all stages of the students to provide intelligent answer services. In this paper, we consider how to make machine solve physical problem in Chinese.

For physical problem in Chinese, its sentence structure and semantics [5] are complex and diversity, this improves the difficulty of problem solving, on the other hand, the machine does not have the congenital ability of understanding and analysing, it cannot correctly identify the semantic information of physical problem. However, it is simple for human. Human can apply unique brain thinking to get the semantic information for problem, and to solve them. So, it is a challenge which makes the machine completely imitate human learning and problem solving.

To make machine think like human in dealing with physical problem, one of the important prerequisites is to let machine understand the semantic information of physical problem and

transform them into a knowledge representation that machine can understand. The study found that knowledge graph [6] can allow machine to understand the semantic information in the physical problem accurately. However the existing knowledge mapping model cannot achieve good results in problem solving [7] because of the lack of expression ability. So this paper presents a model of PKG for physical problem solving, it contains concepts and relations which are split into terms and IDs. From the perspective of knowledge expression, every physical problem is made up of n-tuples which contain concepts and relations between concepts. These n-tuples can be expressed in the form of UP-graph based on PKG. By using UP-graph and the methods stored in the PKG, The physical problem can be solved finally.

The paper presents a novel approach called DKP-solving (solving the physical problem combine deep learning [8, 9] and knowledge graph). It consists of two stages: Firstly, the physical concepts and the relations are extracted by using deep learning, natural language processing (NLP) [10] and term set in PKG and grouped them in the form of n-tuples. By using PKG, UP-graph is generated which is the semantic expression of physical problem. Secondly, the plan of solving problem is given based on the UP-graph and methods stored in PKG. This novel approach of solving the problem is not just for the physical filed, it's also for other discipline, such as mathematic, biology and geography.

The reminder of this paper is organized as follows: Section 2 summarizes the related work about some researches on the automation of problem solving. The proposed novel approach is introduced in Section 3. Section 4 introduces the experiment and the process of physical problem solving. The paper closes with the conclusions and future work in the last section.

2. RELATED WORK

Around the task of problem solving, many companies and laboratories from domestic and foreign have made related studies. Here are some related works of problem solving from some authorities and companies.

In 1964, Bobrow et al. [11] developed the earliest intelligent teaching system-STUDENT which could solve the problem, this system can understand algebraic problem expressed in English, it transforms the text with the form of natural language into a relational model. By manipulating these relational models to achieve the problem of automatic solution. However, STUDENT store sentence structure is very limited.

After the 1980s, with the development of cognitive psychology, especially the classification of problem from semantic layer, promoting the further development of semantic understanding. Riley et al. [12] made a further solving of addition and subtraction problem in 1983, dividing it into three categories: combining problem, transferring problem and comparing problem. And they segmented 14 sub-categories from these three categories. In 1985, Kintsch et al. [13] made a deep cognitive research on one step addition and subtraction application problem. They put forward a characterization model of one step application to solve problem. In 1986, Dellarosa et al.'s [7] research based on the above two theories, the ARITHPRO system was developed which simulates the human cognitive process, to achieve the understanding of one step addition and subtraction application problem.

After the 1990s, the method of statistical natural language processing based on a large corpus were widely used in the field of natural language understanding and achieved good results. Andes Tutorial System is an intelligent tutoring system based on physical mechanics, through understanding the meaning of the problem, completing the interaction with the students [14]and providing problem-solving feedback.

In 2014, Stanford University published Modelling Biological Processes for Reading Comprehension [15] in the EMNLP. Firstly, the biological information process in the text through the forecast to build a rich representation structure, and then use this structure to answer the question, but only for biological knowledge .The Google Deep Learning Group used end-to-end neural networks to solve the problem of reading comprehension, with the application of image-based Attention mechanism in a recurrent neural network (RNN) [16] and the emergence of Attention-based machine translation model [17] and Attention-based image [18] generation. It is usually said that RNN is unable to remember the text of longer distance text in the network related to timing information [19], so Google deep mind mainly uses Attention’s bidirectional Deep Long Short-Term Memory (LSTM) [20], which is an improved version of the bidirectional RNN for solving the reading comprehension. Facebook and other companies and laboratories are also actively studying this issue. Especially Facebook’s babi projects [19] in 2015, Facebook proposed a new Memory Network for understanding problem’ meaning and solving problem according to MeNN [21]. Huawei’s Noah’s Ark Laboratories experimented with Babi’s data set and achieved better results than Facebook’s original test results. They used reasoning model [22] based on GRU [23] neural network. In addition, in 2016, Facebook built another data set for test and training neural network with solving the comprehension reading—children book test, referred to as CBT. In 2016, IBM Watson’s researchers built a new set of depth models-- Attention Sum Reader [24], referred to as ASR which is a model that uses Attention to predict answers directly. This model draws on a structure called Point Network [25]. By summing the Attention values of the same word appearing at different text positions, and the value as the final answer probability distribution value. This model uses the deep learning library theano.

Domestic and foreign related companies and research institutions have made a great contribution for improving the machine ability of problem solving, some researches are in the study of reasoning logical structure, and some using deep learning to solve the problem, but the key point of the present research is to allow machine to understand human intend, and to imitate human behaviour. Chinese sentence structure has unique syntax, complex semantic environment, which deepen the difficulty of problem solving, furthermore some of the knowledge in the application problem is hidden knowledge, more difficult to perceive. Therefore, DKP-solving is proposed to solve the above problems, and the experiment proved that it did improve the machine’s intelligence.

3. DKP-SOLVING ARCHITECTURE

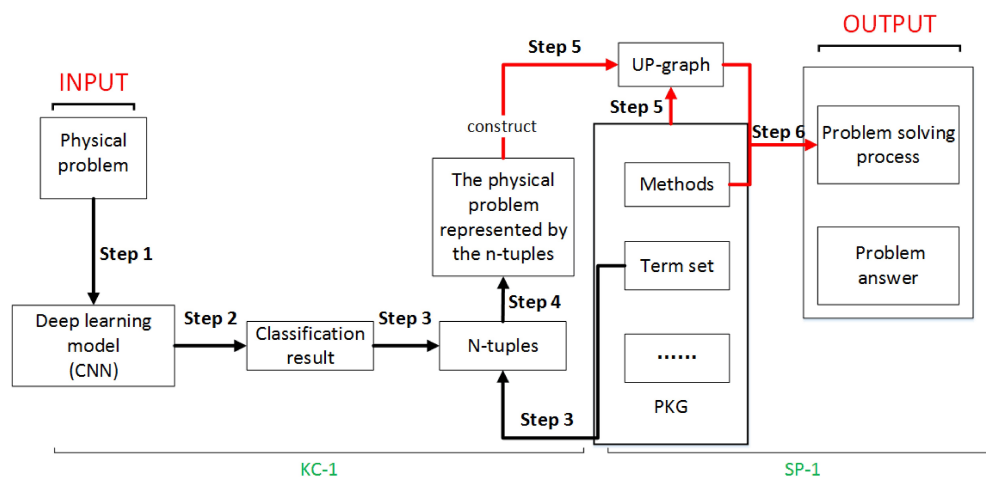


Figure 1. The sketch of DKP-solving

As shown in Figure 1, DKP-solving has two stages. The first stage is called KC-1(First-generation model for extraction of n-tuples) which steps are marked by black arrows, its purpose is to extract n-tuples, which combines the knowledge graph and the Convolutional Neural Network (CNN) [26]. Physical problem (INPUT) is entered the deep learning model (CNN) which built in step1. The classification results for every sentence in physical problem are got by step2. Combining those results with term set in PKG the n-tuple for every sentence are acquired by step3. Finally, physical problem can be represented by the n-tuples in step 4. At the second stage where the steps are marked by red arrows is called SP-1(First-generation for solving problem using knowledge graph), problem's n-tuples will form an UP-graph which is semantic understanding of physical problem in step 5. Under the premise of understanding the meaning of the problem, using the methods stored in the PKG to solve the problem in step 6, and finally get the problem solving process and answer (OUTPUT).

3.1. KC-1's Building

Extracting n-tuples from physical problem helps to express physical semantic information accurately. By matching n-tuples and concepts in PKG, the purpose of building the UP-graph can be achieved which can help solve physical problem.

3.1.1. Definition 1 (Sentence Structure)

By analysing, it is found that the structure of the sentence in physical problem is limited. So the sentences can be divided into 5 types. These five categories are shown in Table 1.

Table 1. Sentence classification

Category	A	B	C	D	E
Sentence Structure	$[h_i_h_i r_i t_i_]$	$[h_i r_i t_i u_i]$	$[_h_i h_i h_i_h_i r_i t_i]$	$[h_i_h_i r_i t_i u_i]$	$[_h_i_h_i r_i h_i_t_i_]$

In Table 1, h_i is a conceptual term in physics, r_i is the relational nouns, characterizing the relations between h_i and t_i , t_i is usually the value of the physical quantity. u_i is the unit of physical concept. “_” is usually a word that does not represent the actual meaning in the sentence and does not affect the overall semantic meaning of the sentence, such as: “此” (this), “的” (of), and “已知” (a known). In order to reasonably explain the meaning of five categories, we give an example, as follows: “小物块的质量为 0.5” (The quality of small block is 0.5), “小物块” (small block) and “质量”(quality) are h_i , “为” (is) is r_i , “0.5” is t_i . It belongs to category A.

3.1.2. Definition 2(Classification-Model)

In order to classify the physical sentence, we use CNN. In physical text analysis, CNN can make it possible to deal with this kind of problem because of the limited length and the compact structure of physical sentence. In this model, the first layer embeds the word into the low-dimensional vector, and the text sequence is expanded into the word vector sequence; The second layer is convoluted with some length different convolution kernels, convoluted followed by the activation function; The third layer is the pooling layer and add dropout; The fourth layer is the fully connected layer and is classified by SoftMax [27].

3.1.3. The Progress of KC-1

Firstly, each sentence in the problem can get the structural classification by the CNN. Secondly, every sentence in physical problem is segmented. Finally, the n-tuples for physical problem are extracted according to the term set of PKG and its struction.

3.2. SP-1's Building

The problem solving process which called SP-1, in this progress, we made a novel knowledge graph model (PKG) which can store semantic information and methods about physical problem.

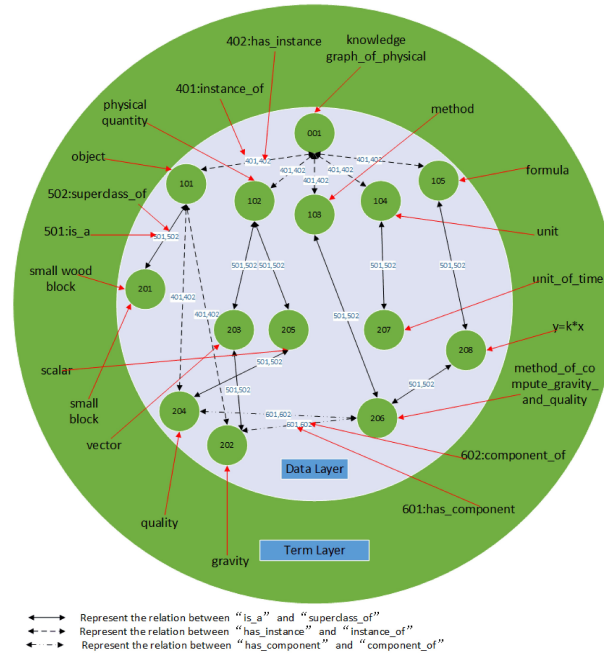


Figure 2. PKG's model structure

Figure.2 is a visual display of PKG's model. The definition of semantic elements based on the PKG model is composed of natural language symbols (terms in the figure) and unique identifier of concepts (represented by ID). Different terms of concepts or relations which have the same meaning use the same ID, such as: "small wood block" and "small block" use the same ID number (201), that is, use the same semantic unit. In the model, Methods are seen as concepts, and have its semantic information and its ID, for example, the method which computes the gravity is split into term "method_of_compute_gravity_and_quality" and ID (206). Furthermore, the method has relations with gravity (202) and quality (204).

3.2.1. The Construction of PKG

Based on the model of PKG, we have built our own PKG which store physical concepts, relations between physical concepts, methods (calculation formulas) in concepts etc. In PKG, we split the physical concept into term and unique ID, term is semantic information of physical concept, and unique ID can be identified by machine, so the representation of physical concepts has two parts, one is term set, and the other is ID set. By this way, machine will be easy to solve the problem about synonymy and polysemy. PKG contains five physical upper concepts, such as "object", "physical quantity", "method", "unit" and "formula", and their unique ID number are "101",

“102”, “103”, “104”, and “105” respectively. There are a number of child nodes below each parent node, such as the parent concept “object” has the child concept “quality” and the child concept “gravity”, in data layer, they are “101”, “204” and “202”.

In PKG, concepts are connected by a certain relation pair which contains two relations, for example, there are two relations (“is_a” and “superclass_of”) between “object” and “small block”, their unique ID number are “501” and “502”. There are two relations (“has_instance” and “instance_of”) between “object” and “quality”, their unique ID number are “402” and “401”, and there are two relations (“has_component” and “component_of”) between “quality” and “method_of_compute_gravity_and_quality”, their unique ID number are “601” and “602”. The concepts and relations formed UP-graph finally. A child concept inherits the nature of the parent concept, for example, child concept “small block” inherits the “quality” and “gravity” of the parent node “object”, unique ID number of four concepts are “201”, “204”, “202” and “101”.

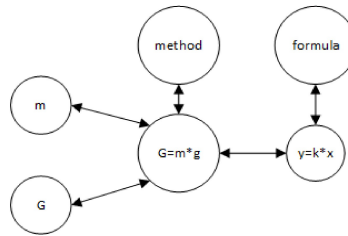


Figure 3. The method in PKG

Method is also a concept and it is the key to solving physical problem, Each method node has its own inputs and outputs, as shown in Figure 3, which provides a method of ($G=m*g$). The method is applied to calculate gravity (G), quality (m), and gravitational acceleration (g). Gravitational acceleration is constant, it is known that any one of gravity (G), quality (m), according the method ($G=m*g$) can be obtained unknown one. Every method has its own calculation form which can be expressed by formula in math, for ($G=m*g$), its form is ($y=k*x$). Table 2 shows some of the methods and the corresponding formula.

Table 2. Methods and formulas

methods	describe	formula
$G=m*g$	Calculate gravity, quality, gravitational acceleration.	$y=k*x$
$m = \rho * v$	Calculate quality, density, volume	
$S=v*t$	Calculate distance, speed, time	
$F=P*S$	Calculate force, pressure, area	
$U=I*R$	Calculate voltage, electric current, resistance	$y=x1*x2$
$I=I1+I2$	Calculate the sum of electric current (series connection)	$y=x1+x2$
$P=\rho gh$	Calculate pressure, density, height, gravitational acceleration	$y=k*x2*x3$

Summary: This graph mainly including the four blocks. There are physical quantities associated with objects, the physical quantity needs to be measured by the unit, solving the problem requires

size are 3×128 , 4×128 , and 5×128 are chosen. For the dropout rate, batch size, learning rate, we set to 0.5, 128, and 0.0001.

4.3. The Process of Solving Physical Problem

As far as we know, DKP is the first approach about solving the physical problems based on knowledge graph and deep learning. This also increases the difficulty of approach comparison, in order to make readers understand DKP better, we used one example to describe it. The result of the calculation is also given at the same time.

Physical problem (example M) is shown below:

一辆大巴车从甲地到乙地运送游客，已知该大巴车在行驶过程中受到的牵引力是重力的0.15倍，质量为5t，时间为11hour，线路的长度为70km，大巴车牵引力做功的功率是多少？(A bus transport tourist from A to B, in the process of this line, traction is 0.15 times that of the gravity, the quality is 5t, time is 11hour, the length of the line is 70km, how much is the power traction work of the bus?)

4.3.1. Extract N-Tuples

The classification is shown in Table 3

Sentence	Category
一辆大巴车从甲地到乙地运送游客, (A bus transport tourists from A to B,)	D
已知该大巴车在行驶过程中受到的牵引力是重力0.15倍, (in the process of this line, traction is 0.15 times that of the gravity,)	E
质量为5t(the quality is 5t,)	B
时间为11hour,(time is 11hour,)	B
线路的长度为70km, (the length of the line is 70km,)	D
大巴车牵引力做功的功率是多少? (how much is the power traction work of the bus?)	C

Table 3. Sentence classification of M

The n-tuples extraction result for M is:

[['大巴车','甲地','乙地'], ['大巴车', '牵引力', '重力', '0.15', '倍'], ['大巴车', '质量', '5', 't'], ['大巴车', '时间', '11', 'hour'], ['线路', '长度', '70', 'km'], ['大巴车', '牵引力', '功率', '? ']] ([['bus', 'A', 'B'], ['bus', 'traction', 'gravity', '0.15', 'times'], ['bus', 'quality', '5', 't'], ['bus', 'time', '11', 'hour'], ['line', 'length', '70', 'km'], ['bus', 'traction', 'power', '? ']])

4.3.2. SP-1's of M

M's UP-graph is formed according to the relations and concepts in Figure 5, "bus" has the attribute of "travel distance", "travel time", "traction", "gravity", "quality". Each attribute has a value and a unit. Such as: "quality" has a value of "5", the unit is "t"(ton). There is a relation between the value of "traction" and "gravity". X is the gravity value. Y is the traction value, Z is the work value, K is the power value. They represent unknowns. This graph lets the machine understand the meaning of each known or unknown physical quantity, and the relations between

them. This part is called UP-graph, the other part stored in the figure are the methods. These methods can be used to solve the M. such as :“method_of_compute_time_power_work”. M’s solving progress are as follows: The value of the “gravity”, “traction” and the “work” is the prerequisite for obtaining the value of “power”.

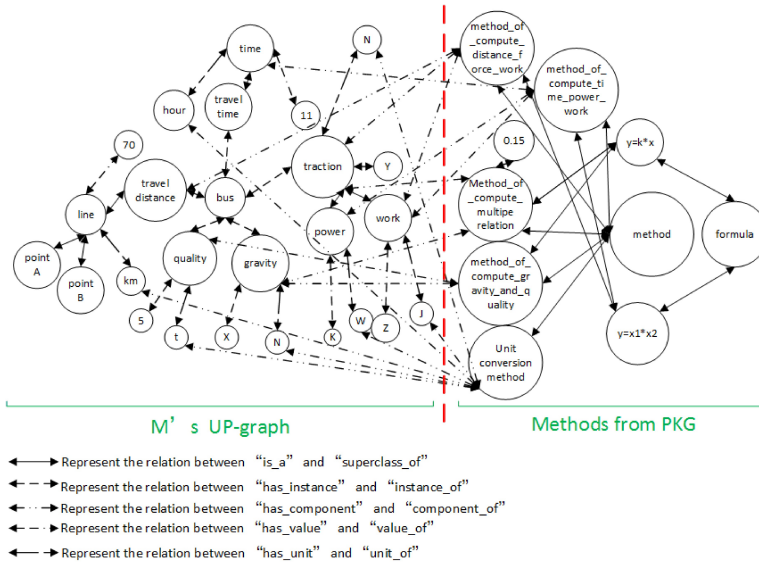


Figure 5. M’s knowledge graph

The machine will use the graph to think like human and use logical reasoning to solve this problem, and it will use the graph to find the question of M is the “power”, the known physical quantities’ value are: “quality”, “line ”, “time ”the unknown physical quantities’ value are: “gravity”, “traction”, “work”, “power” .

M’s solving progresses and result are as follows:

Step 1: Units are normalized according to “Unit conversion method”.

The step (1) for M’s solving progress is shown in Figure 6.

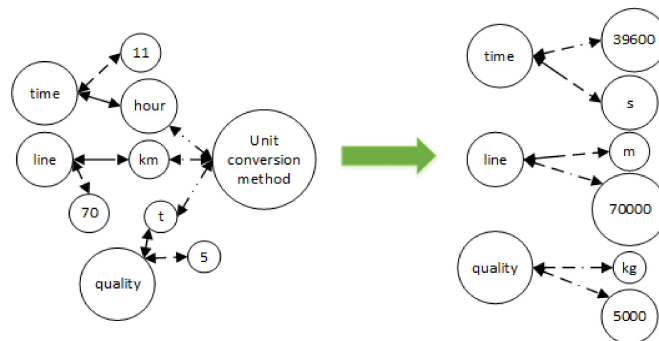


Figure 6. The step (1) for M’s solving progress

By the Figure 6, the value and unit of “time” become “396000” and “s”(second), the value and unit of “line” become “70000” and “m”(meter), the value and unit of “quality” become “5000” and “kg”(kilogram).

Step 2: According to the problem to be solved, we need to find the method of getting the value of “power”. So, the method which named “method_of_compute_time_power_work” is suitable.

Step 3: Find formula in graph which matching the method, and the formula ($y=x1*x2$) is got.

Step 4: Find other physical quantities other than “power” in the method. They are “work” and “time”.

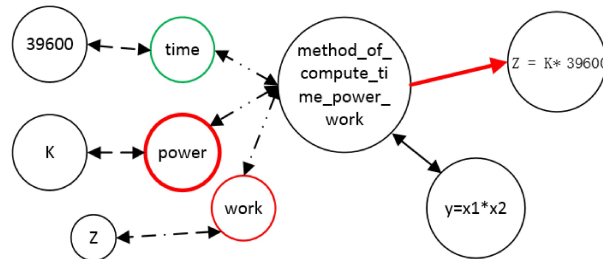


Figure 7. The step (2-4) for M’s solving progress

The step (2-4) for M’s solving progress is shown in Figure 7. Physical quantity in red circle is uncertain, and physical quantity (time) circled by green circle has value (39600). So,

$$K = \frac{Z}{39600} \quad (1)$$

Step 5: Find the method which can get the value of “work”, and the method “method_of_compute_distance_force_work” is suitable.

Step 6: Find the formula which matching this method, the formula ($y=x1*x2$) is got.

Step 7: By the method, “traction” and “line” are found.

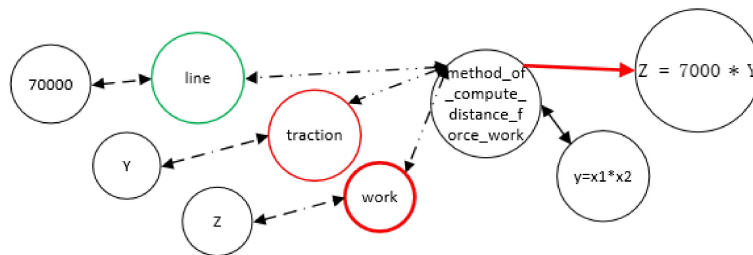


Figure 8. The step (5-7) for M’s solving progress

The step (5-7) for M’s solving progress is shown in Figure 8. Physical quantity in red circle is uncertain, and physical quantity (line) circled by green circle has value (70000). That is,

$$Z = 70000 * Y \quad (2)$$

Step 8: Find the method which can solve the value of “traction”, by the graph, we can find the “gravity” and “traction” has a multiple relation, and the method “Method_of_compute_multiple_relation” can get the value of “traction”.

Step 9: Find the formula to match the method, and the formula is ($y=k*x$).

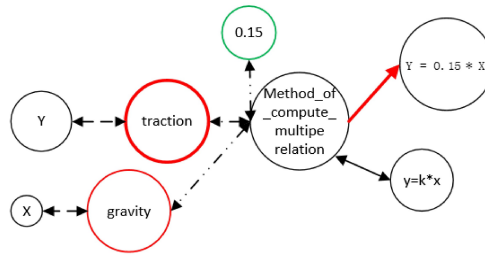


Figure 9. The step (8-9) for M’s solving progress

The step (8-9) for M’s solving progress is shown in Figure 9. Physical quantity in red circle is uncertain, and physical quantity (times) circled by green circle has value (0.15). We can get:

$$Y = 0.15 * X \tag{3}$$

Step 10: Find the quantity which has no value, we get “gravity”.

Step 11: Find the method which can get the value of “gravity”, the method which named “method_of_compute_gravity_and_quality” is got.

Step 12: Find the formula which matching the method, this formula ($y=k*x$) is hit.

Step 13: Find physical quantities other than “gravity” in the method. It is “quality”, and it has value.

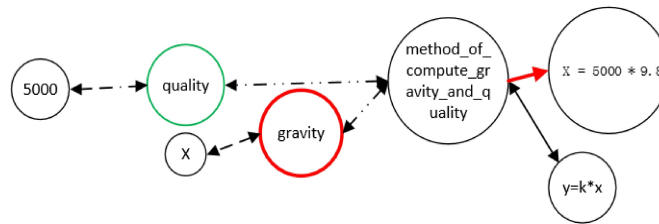


Figure 10. The step (10-13) for M’s solving progress

The step (10-13) for M’s solving progress is shown in Figure 10. Physical quantity in red circle is uncertain, and physical quantity (quality) circled by green circle has value (5000). Gravitational acceleration has default value (9.8). So,

$$X = 5000 * 9.8 \tag{4}$$

Flow chart of M’s solving problem is shown in Figure 11. Finally, the value of “power” and M’s solving progress can be output.

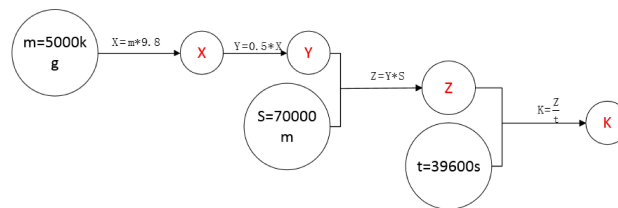


Figure 11. Flow chart of M’s solving problem

$$X = 5000 * 9.8 = 49000 \text{ N} \tag{5}$$

$$Y = 0.15 * X = 0.15 * 49000 = 7350 \text{ N} \quad (6)$$

$$Z = Y * 70000 = 7350 * 70000 = 514500000 \text{ J} \quad (7)$$

$$K = \frac{Z}{39600} = \frac{514500000}{39600} = 12992.42 \text{ w} \quad (8)$$

So, the power of bus traction power is 12992.42 w.

4. CONCLUSIONS

The paper presents a novel approach of solving the physical problem based on Deep Learning and Knowledge Graph: DKP-solving. It combines the CNN and knowledge graph. CNN is used to obtain the sentence structure in physical problem. And then according to the different syntax structure to extract n-tuples. In the process of extracting n-tuples, we are using the classifier and term set of PKG. So every sentence in the problem will be represented by the n-tuples which are used to form the UP-graph. PKG is play an important role in DKP-solving, different from other knowledge graphs, a new model of knowledge graph is constructed which contains concepts, relations, the model can effectively express methods which are key contents in solving problems. By UP-graph and methods, the machine can solve the physical problem effectively and intelligently. In this paper we only use a very traditional way to extract the n-tuples, In the future, we will try more effective n-tuples extraction method, machine solving method, in order to solve the physical problem more effectively. This method also can contribute to other fields, such as Chinese reading Comprehension, mathematical problem solving, and biological computing analysis and so on.

ACKNOWLEDGEMENTS

This research is funded by the Science and Technology Commission of Shanghai Municipality (No. 16511102702) and by Xiaoi Research, by Shanghai Municipal Commission of Economy and Information Under Grant Project No. 201602024 and by the Science and Technology Commission of Shanghai Municipality (No.15PJ1401700).

REFERENCES

- [1] V. Tsiriga and M. Virvou, A Framework for the Initialization of Student Models in Web-based Intelligent Tutoring Systems: Kluwer Academic Publishers, 2004.
- [2] D. Bell, T. Koulouri, S. Lauria, R. D. Macredie, and J. Sutton, "Microblogging as a mechanism for human-robot interaction," Knowledge-Based Systems, vol. 69, pp. 64-77, 2014.
- [3] J. R. Quinlan, C4.5: programs for machine learning: Morgan Kaufmann Publishers Inc., 1993.
- [4] P. Langley, Elements of machine learning: Morgan Kaufmann Publishers Inc., 1995.
- [5] C. Musto, P. Lops, M. D. Gemmis, and G. Semeraro, "Semantics-aware Recommender Systems exploiting Linked Open Data and graph-based features," Knowledge-Based Systems, 2017.
- [6] J. Pujara, H. Miao, L. Getoor, and W. Cohen, "Knowledge Graph Identification," in International Semantic Web Conference, 2013, pp. 542-557.
- [7] D. Dellarosa, "A computer simulation of children's arithmetic word-problem solving," Behavior Research Methods, Instruments, & Computers, vol. 18, pp. 147-154, 1986.

- [8] L. Deng, J. Li, J. Huang, K. Yao, D. Yu, F. Seide, M. Seltzer, G. Zweig, X. He, and J. Williams, "Recent advances in deep learning for speech research at Microsoft," in Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, 2013, pp. 8604-8608.
- [9] T. M. Mitchell, "Does machine learning really work?" AI magazine, vol. 18, p. 11, 1997.
- [10] N. Kalchbrenner, E. Grefenstette and P. Blunsom, "A convolutional neural network for modelling sentences," arXiv preprint arXiv:1404.2188, pp. 655-665, 2014.
- [11] D. G. Bobrow, "Natural Language Input for a Computer Problem Solving System," in Semantic Information Processing, 1964, pp. 281-8.
- [12] M. S. Riley, "Development of children's problem-solving ability in arithmetic.," 1984.
- [13] W. Kintsch and J. G. Greeno, "Understanding and solving word arithmetic problems.," Psychological review, vol. 92, p. 109, 1985.
- [14] R. C. Murray and K. VanLehn, "A comparison of decision-theoretic, fixed-policy and random tutorial action selection," in International Conference on Intelligent Tutoring Systems, 2006, pp. 114-123.
- [15] J. Berant, V. Srikumar, P. Chen, A. Vander Linden, B. Harding, B. Huang, P. Clark, and C. D. Manning, "Modeling Biological Processes for Reading Comprehension.," in EMNLP, 2014.
- [16] V. Mnih, N. Heess and A. Graves, "Recurrent models of visual attention," in Advances in neural information processing systems, 2014, pp. 2204-2212.
- [17] D. Bahdanau, K. Cho and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv:1409.0473, 2014.
- [18] K. Gregor, I. Danihelka, A. Graves, D. J. Rezende, and D. Wierstra, "DRAW: A recurrent neural network for image generation," arXiv preprint arXiv:1502.04623, 2015.
- [19] J. Weston, A. Bordes, S. Chopra, A. M. Rush, B. van Merriënboer, A. Joulin, and T. Mikolov, "Towards ai-complete question answering: A set of prerequisite toy tasks," arXiv preprint arXiv:1502.05698, 2015.
- [20] K. M. Hermann, T. Kocisky, E. Grefenstette, L. Espeholt, W. Kay, M. Suleyman, and P. Blunsom, "Teaching machines to read and comprehend," in Advances in Neural Information Processing Systems, 2015, pp. 1693-1701.
- [21] J. Weston, S. Chopra and A. Bordes, "Memory networks," arXiv preprint arXiv:1410.3916, 2014.
- [22] B. Peng, Z. Lu, H. Li, and K. Wong, "Towards neural network-based reasoning," arXiv preprint arXiv:1508.05508, 2015.
- [23] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arXiv preprint arXiv:1406.1078, pp. 1724-1734, 2014.
- [24] R. Kadlec, M. Schmid, O. Bajgar, and J. Kleindienst, "Text understanding with the attention sum reader network," arXiv preprint arXiv:1603.01547, pp. 908-918, 2016.
- [25] O. Vinyals, M. Fortunato and N. Jaitly, "Pointer networks," in Advances in Neural Information Processing Systems, 2015, pp. 2692-2700.
- [26] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.
- [27] H. B. Demuth, M. H. Beale, O. D. Jess, and M. T. Hagan, Neural Network Design: China Machine Press, 2002.