

Received August 2, 2021, accepted August 22, 2021, date of publication August 30, 2021, date of current version September 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3108768

# Application of Quantum Natural Language Processing for Language Translation

MINA ABBASZADE<sup>1</sup>, VAHID SALARI<sup>2,3,4</sup>, SEYED SHAHIN MOUSAVI<sup>5</sup>, MARIAM ZOMORODI<sup>6,7</sup>, AND XUJUAN ZHOU<sup>8</sup>, (Member, IEEE)

<sup>1</sup>Department of Physics, Isfahan University of Technology, Isfahan 84156-83111, Iran

<sup>2</sup>Basque Center for Applied Mathematics (BCAM), 48009 Bilbao, Spain

<sup>3</sup>Department of Physical Chemistry, University of the Basque Country UPV/EHU, 48080 Bilbao, Spain

<sup>4</sup>Quantum Biology Laboratory, Howard University, Washington, DC 20059, USA

<sup>5</sup>Department of Pure Mathematics, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran

<sup>6</sup>Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad 91779-4897, Iran

<sup>7</sup>Department of Computer Science, Faculty of Computer Science and Telecommunications, Cracow University of Technology, 31-155 Krakow, Poland

<sup>8</sup>School of Business, University of Southern Queensland, Springfield Campus, Springfield Central, QLD 4300, Australia

Corresponding authors: Vahid Salari (vahidsalari742@gmail.com) and Xujuan Zhou (xujuan.zhou@usq.edu.au)

The work of Vahid Salari was supported by the Spanish State Research Agency through BCAM479 Severo Ochoa excellence accreditation SEV-2017-0718, BERC 2018-2021 program, and European projects on Quantum Technologies, QMiCS (Grant No. 820505), EPIQUS (Grant No. 899368), and EU FET Open Grant Quomorphic.

**ABSTRACT** In this paper, we develop compositional vector-based semantics of positive transitive sentences using quantum natural language processing (Q-NLP) to compare the parametrized quantum circuits of two synonymous simple sentences in English and Persian. We propose a protocol based on quantum long short-term memory (Q-LSTM) for Q-NLP to perform various tasks in general but specifically for translating a sentence from English to Persian. Then, we generalize our method to use quantum circuits of sentences as an input for the Q-LSTM cell. This enables us to translate sentences in different languages. Our work paves the way toward representing quantum neural machine translation, which may demonstrate quadratic speedup and converge faster or reaches a better accuracy over classical methods.

**INDEX TERMS** Q-NLP, DisCoCat diagrams, ZX-calculus, quantum circuits, Q-LSTM.

## I. INTRODUCTION

Machine translation (MT) [1] - automated translation of natural languages by computers - was proposed by Warren Weaver in 1949 [2]. Subsequently, a rule-based machine translation (RBMT) was used based on dictionaries and grammars for four decades. Then from the 1980s to 2000s, statistical machine translation (SMT) gained better performance than RBMT and dominated the field. SMT uses statistical models based on the analysis of bilingual text corpus to achieve translation [3]. A few years later in 2003, a language model based on neural networks was suggested [4], which gave a better-quality for the data sparsity problem of traditional SMT models, getting a foundation for the next neural networks on machine translation such as Convolutional Neural Network (CNN), and then Recurrent Neural Network (RNN) as the decoder to transform the state vector into the target language [5]. This resulted in the birth of the Neural Machine Translation (NMT), which is a method that uses deep learning neural networks to map natural

language. NMT's nonlinear mapping differs from the linear SMT models, and describes the semantic equivalence using the state vectors which connect encoder and decoder. NMT is used to predict a sequence of numbers when a sequence of numbers is provided. In the case of translation, each word in the input sentence (e.g English) is encoded as a number to be translated into a resulting sequence of numbers representing the translated target sentence (e.g Persian) via neural network. Vaswani *et al.* introduced the basic transformer encoder-decoder architecture [6]. Martin Popel *et al.* presented a deep learning system, CUBBITT [7], which could win a race (English-to-Czech and Czech-to-English news translation in preserving text meaning) against a professional agency translation in a context-aware blind evaluation by human judge. In general, the success of NMT depends on the quantity and quality of the training pairs of sentences in the source and target language. NMT architecture consists of embedding layers, a classification layer, an encoder network and a decoder network. In the architecture of the CUBBITT machine translation system, the input sentence is converted to a numerical representation that is encoded into a deep representation by a six-layer encoder, which is decoded

The associate editor coordinating the review of this manuscript and approving it for publication was Yuan Zhuang <sup>8</sup>.

by a six-layer decoder into the translation in the target language.

Natural language processing (NLP) [8]–[13] is a subgroup of linguistics and artificial intelligence used for language interactions between computers and humans, e.g. programming computers to analyze natural language data with large volumes. A computer can understand the meanings and concepts of the texts in documents, recognises speech, and generates natural language via NLP. NLP was proposed first in 1950 by Alan Turing [14] i.e. in what is now called the Turing test as a criterion of intelligence for automated interpretation and generation of natural language. Regarding the Turing test for CUBBITT, most participants struggle to distinguish CUBBITT translations from human translations [7]. Also, a group of researchers at OpenAI have developed Generative Pre-trained Transformer 3 (GPT-3) language model [15], as the largest non-sparse language model with higher number of parameters and a higher level of accuracy versus previous models with capacity of ten times larger than that of Microsoft's Turing-NLG to date.

Recent advances in quantum computation and information has opened up new windows in different technologies with broad applications [16]–[21]. For example, recent quantum approaches for NLP have been developed that may reach quantum advantages over classical counterparts in future [22], [23]. Protocols for quantum Natural Language Processing (QNLP) have two aspects: semantic and syntax. Both aspects are performed by a mathematical framework. Compact closed categories are used to provide semantics for quantum protocols [24]. The use of quantum maps for describing meaning in natural language was started by Bob Coecke [25]. Coecke has introduced diagrammatic language to describe processes and their compositions [26]. The diagrammatic language of non-commutative categorical quantum logic represents reduction diagrams for sentences, and allows one to compare the grammatical structures of sentences in different languages. Sadrzadeh has used pregroups to provide an algebraic analysis of Persian sentences [27]. Pregroups are used to encode the grammar of languages. One can fix a set of basic grammatical roles and a partial ordering between them, then freely can generate a pregroup of these types [25]. The category of finite dimensional vector spaces and pregroups are monoidal categories. Models of the semantic of positive and negative transitive sentences are given in ref. [25]. Moreover, Frobenius algebras are used to model the semantics of subject and object relative pronouns [28]. Brian Tyrrell [29] has used vector space distributional compositional categorical models of meaning to compare the meaning of sentences in Irish and in English. Here, we use vector-based models of semantic composition to model the semantics of positive transitive sentences in Persian.

Our goal is to represent an algorithm that can be implemented on quantum hardware and our work is a theoretical proposal in this study. First we need to convert classical information to quantum and vice versa. To go from classical to quantum we employ various encoding quantum circuits.

Parametrized quantum circuits offer a concrete way to implement algorithms and even demonstrate quantum supremacy in the noisy intermediate-scale quantum (NISQ) era. According to [22] the DisCoCat diagram is simplified to some other diagram and is turned into a quantum circuit, which can be compiled via NISQ devices. The grammatical quantum circuits are spanned by a set  $\theta$ . The meaning of the words and hence whole the sentence are encoded in the created semantic space. Finally, we rewrite the diagram as a bipartite graph to turn a quantum circuit. ZX-calculus, like a translator, turns a linguistic diagram into a quantum circuit. According to [30] we consider both grammar and meaning of a grammatical sentence in Persian and turn DisCoCat diagram into a quantum circuit form. In section V we propose an algorithm to translate sentences of different languages. In this algorithm we use the special kind of recurrent neural networks as quantum long short-term memory (LSTM) model. A recurrent neural network (RNN) can be thought of as multiple copies of the same network, each passing a message to a successor. RNNs are not able to learn to connect the information. This problem are solved via the long short-term memory (LSTM). In [31], a hybrid quantum-classical model of LSTM (QLSTM) is proposed. Variational quantum circuits (VQCs) are introduced as the building blocks of the proposed framework. Finally, we present an algorithm that can be used as translating a sentence into its successor. This approach will be applied to convert short English sentences into corresponding Persian sentences. The QLSTM encoder and decoder are used to process the sequence to sequence modelling in this task.

## II. PRELIMINARIES

In this section, we provide some content, which will be used throughout this paper. See the references [28] and [25] for more details.

*Definition 1:* A category  $\mathcal{C}$  consists of:

- a class  $obj(\mathcal{C})$ , called the class of objects;
- for every two objects  $A, B$  a class  $\mathcal{C}(A, B)$  of morphisms; it is convenient to abbreviate  $f \in \mathcal{C}(A, B)$  by  $f : A \rightarrow B$ ;
- for every two morphisms  $f \in \mathcal{C}(A, B)$  and  $g \in \mathcal{C}(B, C)$ , we have  $g \circ f \in \mathcal{C}(A, C)$ . These must satisfy the following properties, for all objects  $A, B, C, D$  and all morphisms  $f \in \mathcal{C}(A, B)$ ,  $g \in \mathcal{C}(B, C)$ ,  $h \in \mathcal{C}(C, D)$ :

$$h \circ (g \circ f) = (h \circ g) \circ f;$$

- for every object  $A$  there is an identity morphism  $1_A \in \mathcal{C}(A, A)$ ; for  $f \in \mathcal{C}(A, B)$  we have

$$1_B \circ f = f = f \circ 1_A$$

*Definition 2:* A monoidal category is a category  $\mathcal{C}$  with the following properties:

- a functor  $\otimes : \mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ , called the tensor product and we have

$$(A \otimes B) \otimes C = A \otimes (B \otimes C);$$

- there is a unit object  $I$  such that

$$I \otimes A = A = A \otimes I;$$

- for each ordered pair morphisms  $f \in \mathcal{C}(A, C), g \in \mathcal{C}(B, D)$  we have  $f \otimes g : A \otimes B \rightarrow C \otimes D$  such that

$$(g_1 \otimes g_2) \circ (f_1 \otimes f_2) = (g_1 \circ f_1) \otimes (g_2 \circ f_2).$$

Monoidal categories are used to encode semantic and syntax of sentences in different languages.

**Definition 3:** A symmetric monoidal category is a monoidal category  $\mathcal{C}$  such that the tensor product is symmetric. This means that there is a natural isomorphism  $\eta$  such that for all objects  $A, B \in \mathcal{C}, A \otimes B \xrightarrow{\eta_{A,B}} B \otimes A$  is an isomorphism.

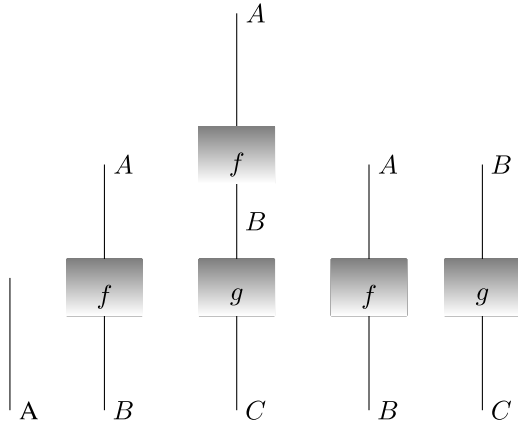
Graphical language is a high-level language for researching in quantum processes, which has applications in many areas such as QNLP and modelling quantum circuits.

### A. GRAPHICAL LANGUAGE FOR MONOIDAL CATEGORY

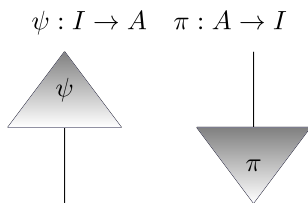
According to [25], morphisms are depicted by boxes, with input and output wires. For example, the morphisms

$$1_A \quad f \quad g \circ f \quad f \otimes g$$

where  $f : A \rightarrow B$  and  $g : B \rightarrow C$ , are depicted as follows:



States and effects of an object  $A$  are defined as follows, respectively from left to right:



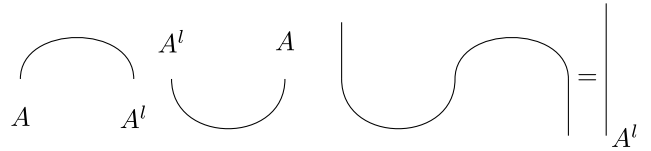
**Definition 4:** A compact closed category is a monoidal category where for each object  $A$  there are objects  $A^r$  and  $A^l$ , and morphisms

$$\begin{aligned} \eta^l : I &\rightarrow A \otimes A^l, & \eta^r : I &\rightarrow A^r \otimes A \\ \epsilon^l : A^l \otimes A &\rightarrow I, & \epsilon^r : A \otimes A^r &\rightarrow I \end{aligned}$$

such that:

- $(1_A \otimes \epsilon^l) \circ (\eta^l \otimes 1_A) = 1_A$
- $(\epsilon^r \otimes 1_A) \circ (1_A \otimes \eta^r) = 1_A$
- $(\epsilon^l \otimes 1_{A^l}) \circ (1_{A^l} \otimes \eta^l) = 1_{A^l}$
- $(1_{A^r} \otimes \epsilon^r) \circ (\eta^r \otimes 1_{A^r}) = 1_{A^r}$ .

The above equations are called yanking equations. In the graphical language the  $\eta$  maps are depicted by caps, and  $\epsilon$  maps are depicted by cups [25]. The yanking equation results in a straight wire. For example, the diagrams for  $\eta^l : I \rightarrow A \otimes A^l, \epsilon^l : A^l \otimes A \rightarrow I$  and  $(\epsilon^l \otimes 1_{A^l}) \circ (1_{A^l} \otimes \eta^l) = 1_{A^l}$  are as follows, respectively from left to right:



**Definition 5:** As defined in [25], a partially ordered non-commutative monoid  $P$  is called a pregroup, to which we refer as *Preg*. Each element  $p \in P$  has both a left adjoint  $p^l \in P$  and a right adjoint  $p^r \in P$ . A partially ordered monoid is a set  $(P, \cdot, 1, \leq, (-)^l, (-)^r)$  with a partial order relation on  $P$  and a binary operation  $\cdot : P \times P \rightarrow P$  that preserves the partial order relation. The multiplication has the unit 1, that is  $p = 1.p = p.1$ . Explicitly we have the following axioms:

$$\begin{aligned} \epsilon_p^l = p^l.p \leq 1, & \quad \epsilon_p^r = p.p^r \leq 1, \\ \eta_p^l = 1 \leq p.p^l, & \quad \eta_p^r = 1 \leq p^r.p \end{aligned}$$

We refer the above axioms as reductions.

### B. PREG AND FVEC AS COMPACT CLOSED CATEGORIES

*Preg* is a compact closed category. Morphisms are reductions and the operation “ $\cdot$ ” is the monoidal tensor of the monoidal category. As mentioned in [25], the category *Preg* can be used to encode the grammatical structure of a sentence in a language. Objects and morphisms are grammatical types and grammatical reductions, respectively. The operation “ $\cdot$ ” is the juxtaposition of types. According to [28], let *FVect* be the category of finite dimensional vector spaces over the field of reals  $\mathbb{R}$ . *FVect* is a monoidal category, in which vector spaces, linear maps and the tensor product are as objects, morphisms and the monoidal tensor, respectively. In this category the tensor product is commutative, i.e.  $V \otimes W \cong W \otimes V$ , and hence  $V^l \cong V^r \cong V^*$ , where  $V^l, V^r$  and  $V^*$  are left adjoint, right adjoint and a dual space of  $V$ . We consider a fixed base, so we have an inner-product. Consider a vector space  $V$  with base  $\{\vec{e}_i\}_i$ . Since  $V$  is an inner product space with finite dimension,  $V^* \cong V$ . Therefore  $V^r \cong V^l \cong V$ ,

$$\begin{aligned} \eta^l = \eta^r : \mathbb{R} &\rightarrow V \otimes V \\ 1 &\mapsto \sum_i \vec{e}_i \otimes \vec{e}_i \end{aligned}$$

and

$$\epsilon^l = \epsilon^r : V \otimes V \rightarrow \mathbb{R} \sum_{ij} c_{ij} \vec{v}_i \otimes \vec{w}_j \mapsto \sum_{ij} c_{ij} \langle \vec{v}_i | \vec{w}_j \rangle.$$

**TABLE 1.** Examples of the equivalent expressions in English and Persian.

English	Persian	فارسی
Sara	Sara	سارا
buys	mikharad	می‌خرد
the	ra	را
book	ketab	کتاب
Sara buys the book	Sara ketab ra mikharad	سارا کتاب را می‌خرد

Consider the monoidal functor  $F : Preg \rightarrow FVect$ , which assigns the basic types to vector spaces as follows:

$$F(n) = N \quad F(s) = S \quad F(1) = I,$$

and also  $F(x \otimes y) = F(x) \otimes F(y)$ . The compact structure is preserved by Monoidal functors; this means that

$$F(x^r) = F(x^l) = F(x)^*$$

for more details see [28].

### III. POSITIVE TRANSITIVE SENTENCE

The simple declarative Persian sentence with a transitive verb has the following structure: subject + object + objective sign + transitive verb. For example, ‘Sara ketab ra mikharad’ is the Persian sentence for ‘Sara buys the book’. In Table 1, we present the equivalent expressions in English and Persian. In this sentence, ‘Sara’ is the subject, ‘ketab’ is the direct object, ‘ra’ is the objective sign and ‘mikharad’ is the transitive verb in present tense, see [27].

#### A. VECTOR SPACE INTERPRETATION

Vector spaces and pregroups are used to assign meanings to words and grammatical structure to sentences in a language. The reductions and types are interpreted as linear maps and vector spaces, obtained by a monoidal functor  $F$  from  $Preg$  to  $FVect$ . In this paper we present one example from Persian: positive transitive sentence, for which we fix the following basic types,

- n: noun
- s: declarative statement
- o: object

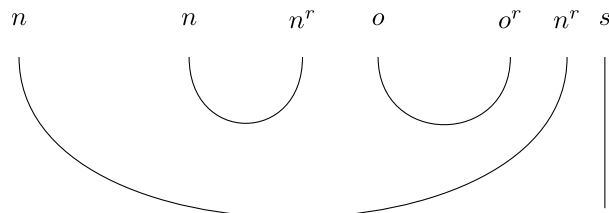
According to [25] if the juxtaposition of the types of the words in a sentence reduces to the basic type  $s$ , the sentence is called grammatical. We use an arrow  $\rightarrow$  for  $\leq$  and drop the “.” between juxtaposed types. According to [27] the example sentence ‘Sara ketab ra mikharad’, has the following type assignment:

$$\begin{matrix} \text{Sara} & \text{ketab} & \text{ra} & \text{mikharad.} \\ n & n & (n^r o) & (o^r n^r s) \end{matrix}$$

Which is grammatical because of the following reduction:

$$nn(n^r o)(o^r n^r s) \rightarrow n11n^r s \rightarrow s$$

For the above sentence this reduction is depicted diagrammatically as follows:



**FIGURE 1.** The reduction diagram of the sentence.

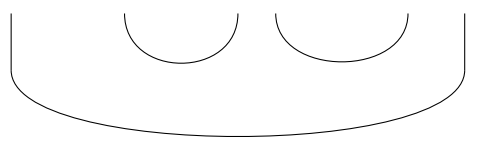
A positive sentence with a transitive verb in Persian has the pregroup type  $nn(n^r o)(o^r n^r s)$ . The interpretation of a transitive verb is computed as follows:

$$\begin{aligned} F(o^r \otimes n^r \otimes s) &= F(o^r) \otimes F(n^r) \otimes F(s) \\ &= F(o)^r \otimes F(n)^r \otimes F(s) \\ &= F(o)^* \otimes F(n)^* \otimes F(s) \\ &= N \otimes N \otimes S. \end{aligned}$$

So the meaning vector of a Persian transitive verb is a vector in  $N \otimes N \otimes S$ . The pregroup reduction of a transitive sentence is computed as follows:

$$\begin{aligned} F((\epsilon_n^r \otimes 1_s) \circ (1_n \otimes \epsilon_n^r \otimes \epsilon_o^r \otimes 1_n \otimes 1_s)) \\ &= (F(\epsilon_n^r) \otimes F(1_s)) \circ (F(1_n) \otimes F(\epsilon_n^r) \otimes F(\epsilon_o^r) \otimes F(1_s)) \\ &= (F(\epsilon_n)^* \otimes F(1_s)) \circ (F(1_n) \otimes F(\epsilon_n)^* \otimes F(\epsilon_o)^* \otimes F(1_s)) \\ &= (\epsilon_N \otimes 1_S) \circ (1_N \otimes \epsilon_N \otimes \epsilon_N \otimes 1_N \otimes 1_S) \end{aligned}$$

and depicted as:



The distributional meaning of ‘Sara ketab ra mikharad’ is as follows:

$$\begin{aligned} F((\epsilon_n^r \otimes 1_s) \circ (1_n \otimes \epsilon_n^r \otimes \epsilon_o^r \otimes 1_n \otimes 1_s)) \\ \overrightarrow{(Sara \otimes ketab \otimes ra \otimes mikharad)} \end{aligned}$$

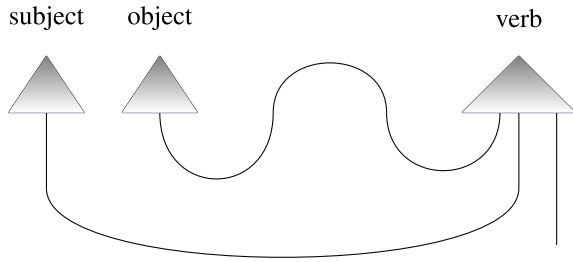
where  $\overrightarrow{ra}$  is the vector corresponding to the meaning of ‘ra’. We set

$$\overrightarrow{ra} = \sum_i \overrightarrow{e}_i \otimes \overrightarrow{e}_i \in N \otimes N$$

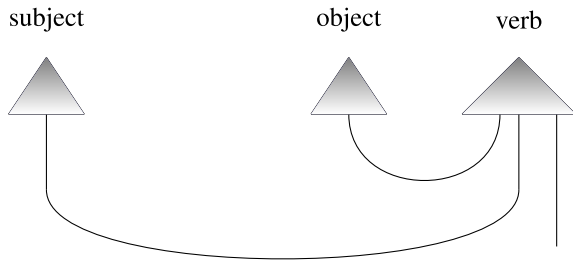
and in this case we have

$$\vec{ra} \simeq \eta_N : \mathbf{R} \rightarrow N \otimes N.$$

We obtain diagrammatically:



Which by the diagrammatic calculus of compact closed categories [32], is equal to:



Consider the vector  $\Psi$  in the tensor space which represents the type of verb:

$$\Psi = \sum_{ijk} c_{ijk} \vec{w}_i \otimes \vec{v}_j \otimes \vec{s}_k \in N \otimes N \otimes S$$

where for each  $i$ ,  $\vec{w}_i$  is the meaning vector of object and  $\vec{v}_j$  is the meaning vector of subject. Then we have:

$$\begin{aligned} & (\epsilon_N \otimes 1_S) \circ (1_N \otimes \epsilon_N \otimes 1_N \otimes 1_S) (\vec{v} \otimes \vec{w} \otimes \vec{\Psi}) \\ &= (\epsilon_N \otimes 1_S) \circ (1_N \otimes \epsilon_N \otimes 1_N \otimes 1_S) \\ & (\vec{v} \otimes \vec{w} \otimes (\sum_{ijk} c_{ijk} \vec{w}_i \otimes \vec{v}_j \otimes \vec{s}_k)) \\ &= (\epsilon_N \otimes 1_S) (\vec{v} \otimes \langle w, w_i \rangle \otimes \vec{v}_j \otimes \vec{s}_k) \\ &= \sum_{ijk} c_{ijk} \langle w, w_i \rangle \langle v, v_j \rangle \vec{s}_k. \end{aligned}$$

### B. TRUTH THEORETIC MEANING AND CONCRETE INSTANTIATION

According to [28] we let  $N$  to be the vector space spanned by a set of individuals  $\{\vec{n}_i\}$  and  $S$  to be the one dimensional space spanned by the unit vector  $\vec{1}$ . The unit vector and the zero vector represent truth value 1 and truth value 0 respectively. A transitive verb  $\Psi \in N \otimes N \otimes S$  is represented as follows:

$$\Psi := \sum_{ji} \vec{n}_j \otimes \vec{n}_i \otimes (\alpha_{ji} \vec{1})$$

where  $\vec{sub} = \sum_i \vec{n}_i$ ,  $\vec{obj} = \sum_j \vec{n}_j$  and  $\alpha_{ji}$ 's are degrees of truth, i.e.  $\vec{n}_i$   $\Psi$ 's  $\vec{n}_j$  with degree  $\alpha_{ji}$ , for all  $i, j$ . For

$\vec{sub} = \sum_k \vec{n}_k$  and  $\vec{obj} = \sum_l \vec{n}_l$ , where  $k$  and  $l$  range over the sets of basis vectors representing the respective common nouns, the truth-theoretic meaning of a transitive sentence is computed as follows:

$$\begin{aligned} & \overrightarrow{sub \ obj \ ra \ verb} \\ &= (\epsilon_N \otimes 1_S) \circ (1_N \otimes \epsilon_N \otimes 1_N \otimes 1_S) (\vec{sub} \otimes \vec{obj} \otimes \vec{verb}) \\ &= (\epsilon_N \otimes 1_S) \circ (1_N \otimes \epsilon_N \otimes 1_N \otimes 1_S) (\sum_k \vec{n}_k \otimes \sum_l \vec{n}_l \otimes \\ & (\sum_{ji} \vec{n}_j \otimes \vec{n}_i \otimes \alpha_{ji} \vec{1})) = \sum_{kl} \alpha_{kl} \vec{1}. \end{aligned}$$

For concrete instantiation in the model of Grefenstette and Sadrzadeh [33] the vectors are obtained from corpora and the scalar weights for noun vectors are not necessarily 1 or 0. For any word vector  $\vec{word} = \sum c_i^{word} \vec{n}_i$ , the scalar weight  $c_i^{word}$  is the number of times that the word has appeared in that context. Where  $\vec{n}_i$ 's are context basis vectors. The meaning of the transitive sentence is:

$$\overrightarrow{sub \ obj \ ra \ verb} = \sum_{jit} \langle \vec{obj} | \vec{n}_j \rangle \langle \vec{sub} | \vec{n}_i \rangle c_{jit} \vec{s}_t$$

A transitive verb is represented as a two dimensional matrix. The corresponding vector of this matrix is  $\vec{verb} = \sum_{ji} c_{ji} (\vec{n}_j \otimes \vec{n}_i)$ . Note that the sum of the tensor product of the objects and subjects of the verb throughout a corpus represents the meaning vector of the verb. So the meaning of the transitive sentence is:

$$\begin{aligned} & \overrightarrow{sub \ obj \ ra \ verb} = \sum_{ji} \langle \vec{obj} | \vec{n}_j \rangle \langle \vec{sub} | \vec{n}_i \rangle c_{ji} (\vec{n}_j \otimes \vec{n}_i) \\ &= \sum_{ji} c_j^{obj} c_i^{sub} c_{ji} (\vec{n}_j \otimes \vec{n}_i). \end{aligned}$$

The meaning vector is decomposed to point-wise multiplication of two vectors as follows:

$$\begin{aligned} & (\sum_{ji} c_j^{obj} c_i^{sub} (\vec{n}_j \otimes \vec{n}_i)) \odot (\sum_{ji} c_{ji} (\vec{n}_j \otimes \vec{n}_i)) \\ &= (\vec{obj} \otimes \vec{sub}) \odot \vec{verb} \end{aligned}$$

where  $\odot$  is the point-wise multiplication. The meaning vector of the transitive sentence in English is as follows

$$(\vec{sub} \otimes \vec{obj}) \odot \vec{verb}.$$

Thus for synonymous sentences in the same corpus of English and Persian we have:

$$(\vec{sub} \otimes \vec{obj}) \odot \vec{verb} = (\vec{obj} \otimes \vec{sub})^T \odot \vec{verb}$$

where  $(\vec{obj} \otimes \vec{sub})^T$  is the transpose of  $(\vec{obj} \otimes \vec{sub})$ . The procedure for learning weights for matrices of subject, object and verb is represented in [33]. Note that  $(\vec{obj} \otimes \vec{sub})$  is the Kronecker product of  $\vec{obj}$  and  $\vec{sub}$  of the verb. In order to compare the meaning of two sentences in English and Persian, we compute the following:

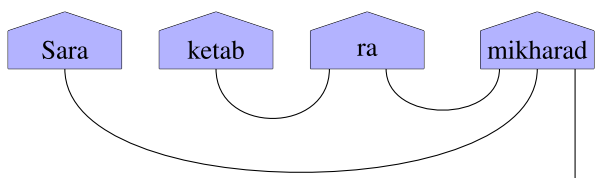
$$(\vec{sub} \otimes \vec{obj}) \odot \vec{verb} - (\vec{obj} \otimes \vec{sub})^T \odot \vec{verb}.$$



Our goal in classification synonymous sentences in two different languages is to translate the sentence from one language to another.

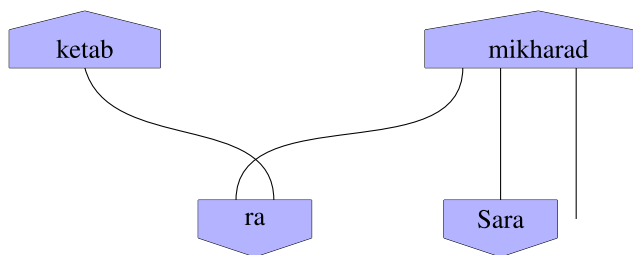
**IV. DIAGRAMS REWRITING AND QUANTUM CIRCUITS**

As mentioned in the previous sections a sentence in a corpus is parsed according to its grammatical structure. According to [22] we simplify the DisCoCat diagram to some other diagram and turn into a quantum circuit, which can be compiled via NISQ devices. Two methods are presented for this purpose. The bigraph method and snake removal method. Both methods are done in the symmetric version of the pregroup grammar. We consider the grammatical sentence from Fig. 1, and obtain the diagram in Fig. 2.



**FIGURE 2.** The DisCoCat diagram of the sentence.

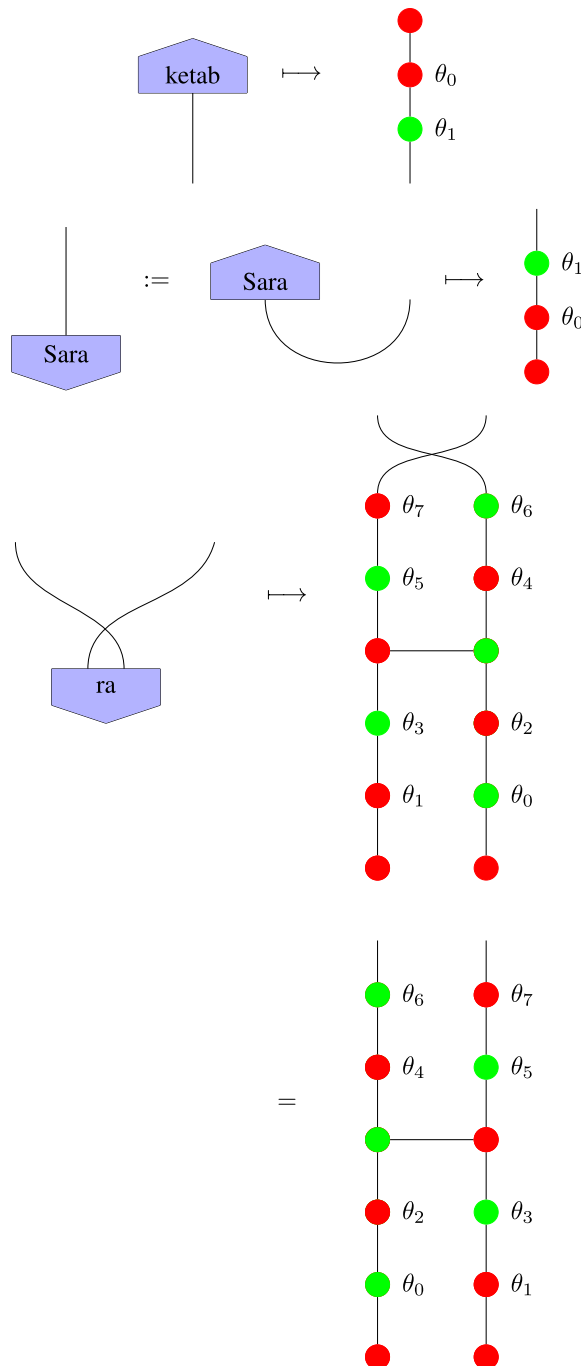
where the boxes denote tensors, and the order of which is the number of their wires. We use a bigraph method to turn the diagram in Fig. 2 into a bipartite graph. Words at odd distance from the root word are transposed into effects, and we obtain the diagram in Fig. 3.



**FIGURE 3.** The rewritten diagram.

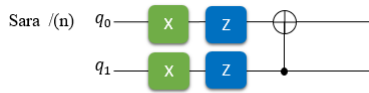
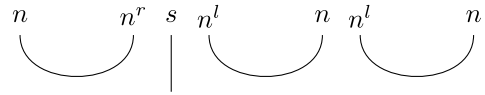
Transposition turns states into effects, see [32]. According to [22], we consider CNOT+U(3) of unitary qubit ansatz. Layers of CNOT gates between adjacent qubits with layers of single-qubits rotations in Z and X form unitary quantum circuits. Let  $\hat{P}$  be the symmetric version of the pregroup grammar  $P$ . Consider the monoidal functor from  $\hat{P}$  to  $\mathbb{fHilb}$ , in which word states are mapped to the state ansatz. State ansatz are obtained by applying the unitary ansatz to the pauli  $Z|0\rangle$  state. Word effects are mapped to the effect ansatz, in which effect ansatz are obtained by transposing the state ansatz in the computational basis, and wire crossing are mapped to swaps. Now consider the diagram in Fig. 3. If each wire is mapped to a qubit, the circuit has about four CNOTs. In ZX-calculus [34], suppose single-qubits white and black dots are rotations in pauli Z and pauli X. A CNOT

gate is Black and white dots connected by a horizontal line. So we obtain the following circuits:

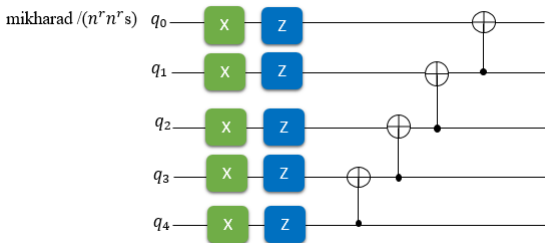
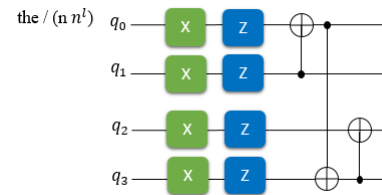
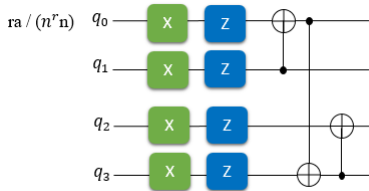
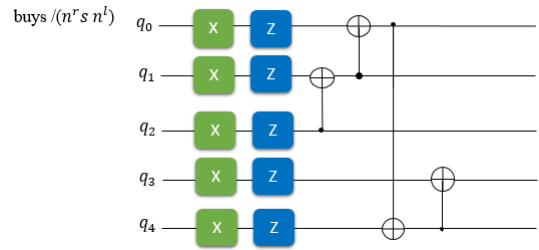


One can use the bigraph algorithm to form quantum circuits of the semantic side of the meaning. In the pregroup type of the sentence ‘Sara ketab ra mikharad’ set  $o = n$ . For atomic types  $n$  and  $s$  consider two qubits and one qubit respectively. The number of qubits for each type  $t$  is the sum of the number of qubits associated to all atomic types in  $t$ . For example the transitive verb ‘mikharad’ has five qubits.

For each word in the sentence we have a quantum circuit as follows:



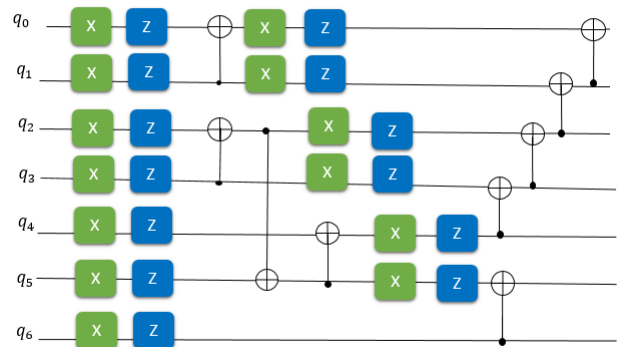
and the quantum circuit of each word is as follows:



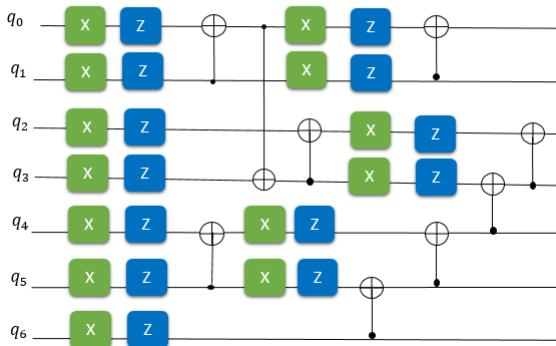
So the quantum circuit of the whole sentence is as follows:

The quantum circuit of the whole sentence is as follows:

Sara buys the book



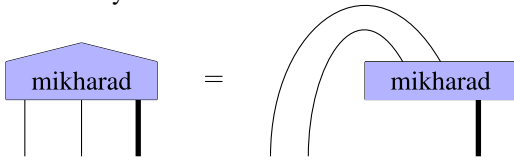
Sara kitab ra mikharad



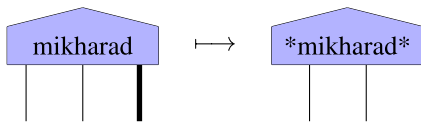
The reduction diagram of the sentence ‘Sara buys the book’ in English is:

The two sentences ‘Sara Kitab ra mikharad’ and ‘Sara buys the book’ have the same meaning but are grammatically different. By training these circuits on a quantum hardware, we’ll give a model to translate from English to Persian and vice versa. According to [30] we present grammar+meaning as quantum circuit for the above two sentences. Consider the states  $|\psi_{n_s}\rangle$  and  $|\psi_{n_o}\rangle$  correspond to the subject and the object,

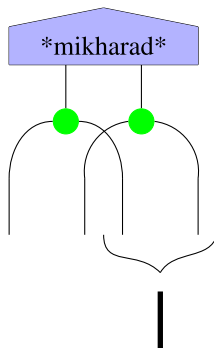
respectively. Also a transitive verb as a map  $\eta_{tv}$  that takes  $|\psi_{n_s}\rangle \in \mathbb{C}^2$  and  $|\psi_{n_o}\rangle \in \mathbb{C}^2$  and produces  $|\psi_{n_s.n_o.tv}\rangle \in \mathbb{C}^{2^k}$ , diagrammatically:



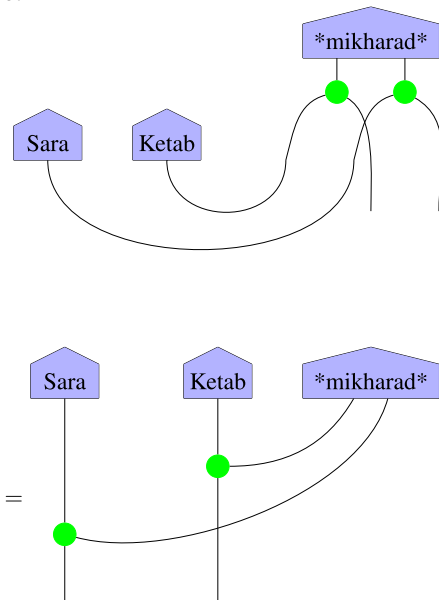
So  $|\psi_{mikharad}\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^{2^k}$ . Because the quantum model relies on the tensor product, an exponential blow-up occurs for meaning spaces of words. In order to avoid this obstacle in experiments decrease the dimension of the spaces in which meanings of transitive verbs live. For the transitive verb, instead of state in the large space  $|\psi_{mikharad}\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^{2^k}$  consider state in a smaller space  $|\psi_{*mikharad*}\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ , diagrammatically:



Then copy each of the wires and bundle two of the wires together to make up the thick wire. Thus 'mikharad' is obtained:



For more details see [30]. Now inter Sara and Ketab into the picture:



we pull some spiders out and obtain:

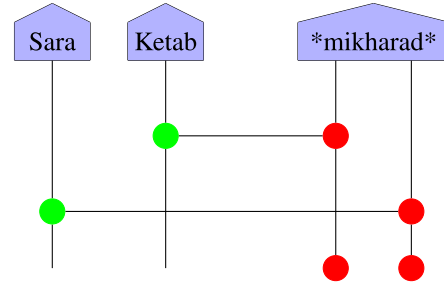


FIGURE 4. The circuit obtained from pulling out spiders.

and by Using the Choi-Jamiolkowski correspondence we obtain the circuit in Fig. 5. The circuit in Fig. 4 requires 4 qubits and has two CNOT-gates in parallel, but the circuit Fig. 5 requires 3 qubits and has sequential CNOT-gates. Indeed, the use of the Choi-Jamiolkowski correspondence has reduced the number of qubits, but has increased the depth of the CNOT-gates.

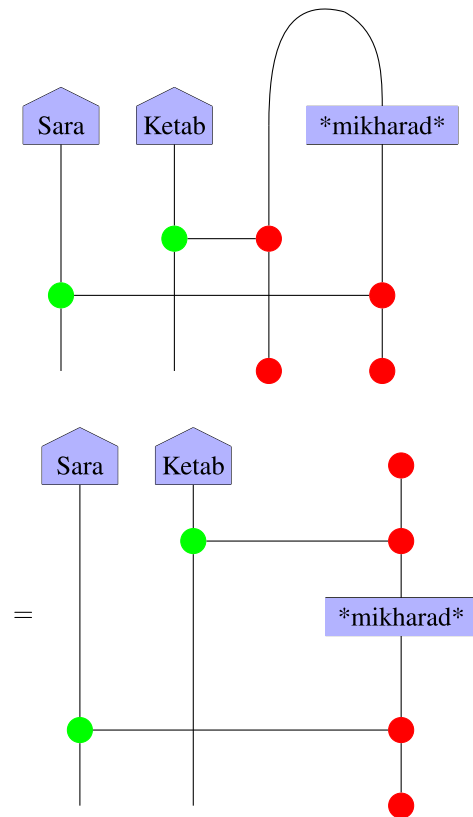


FIGURE 5. The circuit obtained from Choi-Jamiolkowski correspondence.

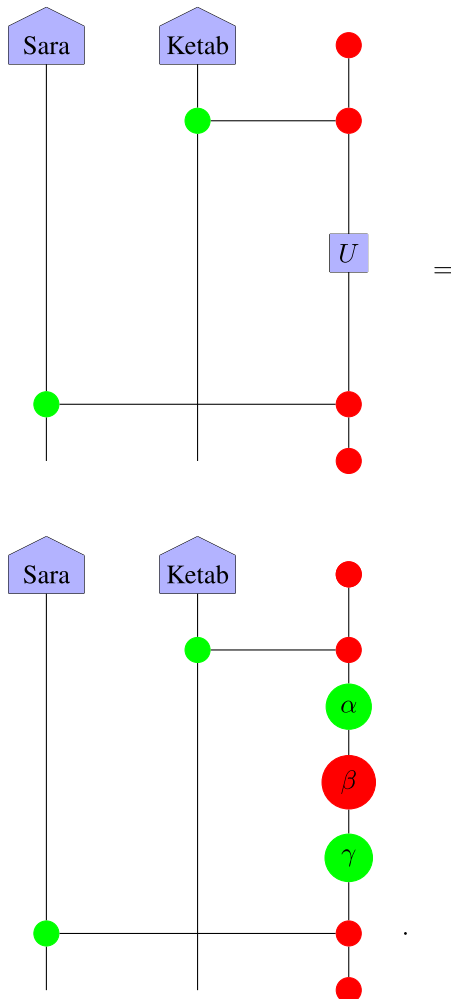
As mentioned in [30] ion trap hardware has less qubits, but performs better for greater circuit depth. In ZX-calculus and via Euler decomposition any one-qubit unitary gate is



represented as follows:



Each verb is represented by an unitary gate  $U$  and has different values  $\alpha, \beta$  and  $\gamma$ . So we obtain the following circuits:



By considering the singular value decomposition for the verb we obtain the circuit of Fig. 6. Where the state  $P$  is the diagonal of the matrix. We represent all noun states by gates and obtain the circuit of Fig. 8. As mentioned in III-A, similarly for the sentence ‘Sara buys the book’ we obtain the DisCoCat diagram in Fig. 7. Indeed we ignore ‘the’ and ‘ra’ of positive transitive sentences in English and Persian respectively. Therefore according to [30] the parametrised quantum circuit of the diagram in Fig. 7 is as Fig. 9. The

quantum circuits in Fig. 8 and Fig. 9 are compatible to train on a quantum hardware. By providing a set of quantum circuits of sentences in English and Persian and training on a quantum hardware we can perform the classification task. This approach can make a quicker and an advantage over classical counterparts for NLP. In the next section we try to provide an algorithm for translating sentences of different languages.

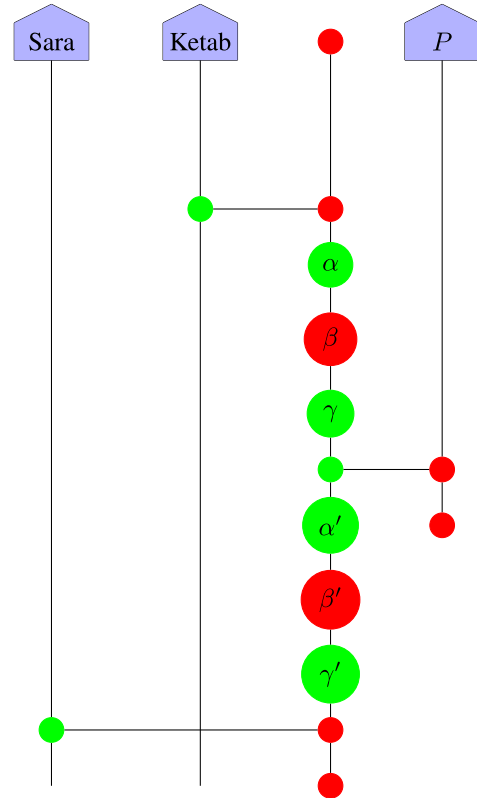


FIGURE 6. The circuit-form of ‘Sara ketab ra mikharad’.

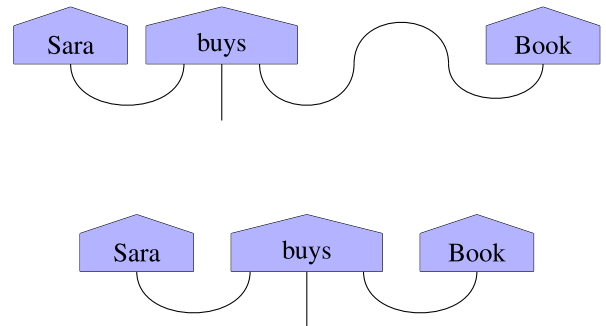


FIGURE 7. DisCoCat diagram of the sentence ‘Sara buys the book’.

### V. ALGORITHM OF USING QUANTUM LONG SHORT-TERM MEMORY FOR QUANTUM NATURAL LANGUAGE PROCESSING

In this section we propose an algorithm to translate sentences of different languages. In this algorithm we use quantum

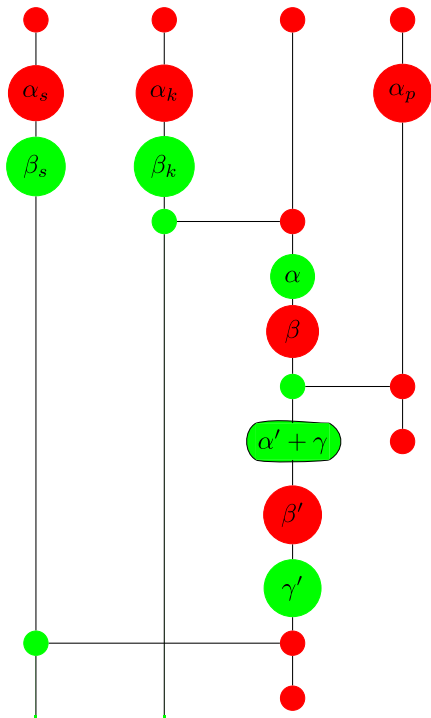


FIGURE 8. The quantum circuit of 'Sara ketab ra mikharad'.

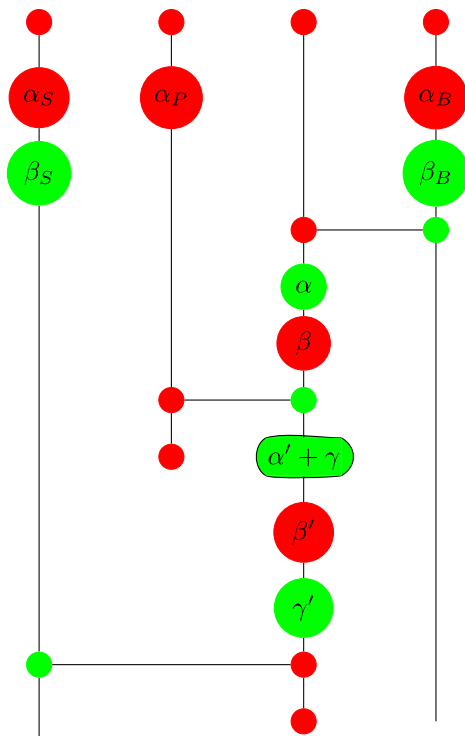


FIGURE 9. The quantum circuit of 'Sara buys the book'.

long short-term memory (LSTM) model. The long short-term memory is a special kind of recurrent neural networks (RNNs) that can learn a longer range of sequential dependency in the data. In this way, we use parametrized quantum circuits (PQCs) of sentences as quantum input for

quantum LSTM cell. This model encodes features of a quantum circuit of a sentence into a learned hidden vector state and then decodes that vector into its successor sentence. A hybrid quantum model of LSTM to translate is capable of modeling sequential data. Now we use the method presented in [35] to map a DisCoCat diagram into a quantum circuit. Sentences of different grammatical structure are mapped to different quantum circuits. This method is easier than the methods presented in section IV. In DisCoCat diagram of the sentence 'Sara ketab ra mikharad' in Fig. 2 we bend down all nouns of the sentence, and obtain the diagram of Fig. 10. Then we have the quantum circuit of Fig. 11.

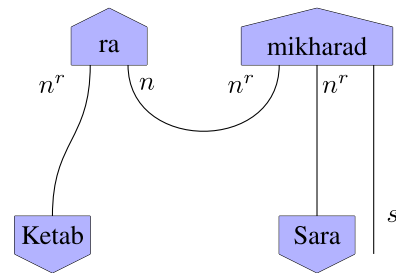


FIGURE 10. The rewritten diagram.

We choose one qubit for every wire of type  $n$  and  $s$ , and replace all word states (effects) with parametrized quantum states (effects). So the verb is a state on three qubits. For verb, so called IQP-based states is being used. IQP layer consists of an H gate on every qubit composed with two controlled Z-rotation gates, connecting adjacent qubits. The words 'Sara' and 'ketab' replace with the quantum effects. Note that each cup is corresponding to a Bell-effect. Similarly for the DisCoCat diagram of the sentence 'Sara buys the book' we have the diagram of Fig. 12, and obtain the quantum circuit of Fig.13.

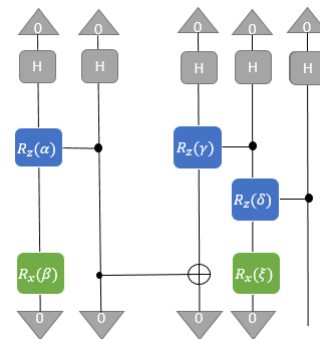


FIGURE 11. The quantum circuit of 'Sara ketab ra mikharad'.

Consider a language model trying to translate a sentence. Encoder refers to the part of the network which reads the sentence to be translated, and decoder is the part of the network which translates the sentence into desired language. The first step of our encoding scheme is to transform sentences into PQCs. We choose features of the circuit of each sentence as

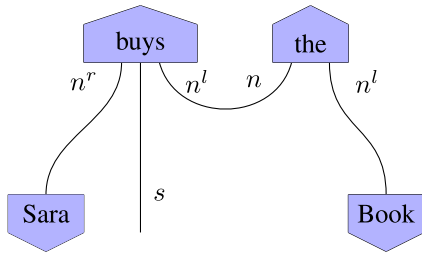


FIGURE 12. The rewritten diagram.

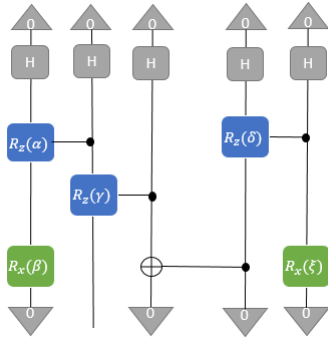


FIGURE 13. The quantum circuit of ‘Sara buys the book.’

an input vector of a quantum LSTM cell. Quantum circuits of sentences are trained on a quantum processor. At each time step  $t$ , the aforementioned input is composed of a set of features characterizing both the circuit and the processor. In particular, with respect to the circuit, the following information can be considered:

- an integer value representing the number of qubits composing the circuit;
- an integer value representing the total number of CNOT gates in the circuit;
- an integer value representing the total number of Z-rotation gates in the circuit;
- an integer value representing the number of quantum effects in the circuit;
- a matrix of integer values where the item  $[i, j]$  contains the number of CNOT gates between the control qubit and the target qubit of the circuit.

The task we present can be thought as translating a sentence into its successor. We propose an algorithm of using quantum LSTM for quantum natural language processing in order to perform various tasks such as prediction a sentence and translating. The algorithm in order to translate a sentence from English to Persian is pictured diagrammatically in Fig. 14. The encoder-decoder model of quantum LSTM (QLSTM) encodes a feature vector of the quantum circuit of the sentence into a learned hidden vector state and then decodes that into its successor sentence. The QLSTM cell of the algorithm can be considered as shown in Fig. 15, [31]. In the QLSTM cell, there are six VQCs. For VQC<sub>1</sub> to VQC<sub>4</sub> consider the concatenation  $v_t$  of the hidden state  $h_{t-1}$  from the previous time step and current input vector  $x_t$ . We choose

features of the quantum circuit of each sentence as input vector  $x_t$ . The measured values of VQCs go through nonlinear activation functions. A mathematical formulation of the QLSTM cell is as follows:

$$\begin{aligned}
 f_t &= \sigma(\text{VQC}_1(v_t)), \quad v_t = [h_{t-1}, x_t] \\
 i_t &= \sigma(\text{VQC}_2(v_t)) \\
 \tilde{C}_t &= \tanh(\text{VQC}_3(v_t)) \\
 c_t &= f_t * c_{t-1} + i_t * \tilde{C}_t \\
 o_t &= \sigma(\text{VQC}_4(v_t)) \\
 h_t &= \text{VQC}_5(o_t * \tanh(c_t)) \\
 y_t &= \text{VQC}_6(o_t * \tanh(c_t))
 \end{aligned}$$

An output of a vector  $f_t$  has values in  $[0,1]$  through the sigmoid function.  $*$  is the element-wise multiplication.  $f_t * c_{t-1}$  determines whether to keep or forget the elements in the cell state  $c_{t-1}$  from the previous step. The output of the VQC<sub>2</sub> goes through the sigmoid function to determine which values will be added to the cell state. The output of the VQC<sub>3</sub> goes through the tanh function to generate a new cell state  $\tilde{C}_t$ .  $i_t * \tilde{C}_t$  is used to update the cell state. The output of the VQC<sub>4</sub> goes through the sigmoid function, and determines which values in the cell state  $c_t$  are relevant to the output.  $o_t * \tanh(c_t)$  is processed with VQC<sub>5</sub> to get the hidden state  $h_t$ . Also  $o_t * \tanh(c_t)$  is processed with VQC<sub>6</sub> to get the output  $y_t$ . Generic VQC architecture is shown in Figure 16. In the VQC architecture,  $x_i$ 's are elements of the vector  $v_t$ . The number of qubits and the number of measurements are determined by the numbers of  $x_i$ 's. Three rotation angles  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$  are not fixed. They are updated in the iterative optimization process [31].

In the algorithm presented, one may face some problems.

The first three sentences in Table 2 have the same structure but different meanings, with quantum circuits such as in Fig. 13. Each sentence may have different values for  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\xi$ . For entering this feature to the algorithm, one can enter this angles as real numbers into the input vectors. In Table 2, the last two sentences have the same meanings but different structures. For each language, one has to do the meaning classification task in QNLP. As mentioned in the III-A, the meaning of sentences can be reduced to linear algebraic formulae. For example, the meaning vector of our transitive sentence is:

$$\overrightarrow{\text{Sara buys the book}} = f(\overrightarrow{\text{Sara}} \otimes \overrightarrow{\text{buys}} \otimes \overrightarrow{\text{the}} \otimes \overrightarrow{\text{book}})$$

where  $f$  is the linear map that encodes the grammatical structure. Learning a semantic vector for each word is learning it's basis weights from the corpus. This setting offers geometric means to reason about semantic similarity, e.g. via Cosine measure. Grefenstette et al. [36] give some similarity calculations for some sentence pairs via inner product between the meaning vector of sentences. In Table 3, one may translate the first three sentences in Table 2 to Ukrainian and Spanish, and can consider grammatical sentences of any language, and determine the type of words in the sentences. One can verify

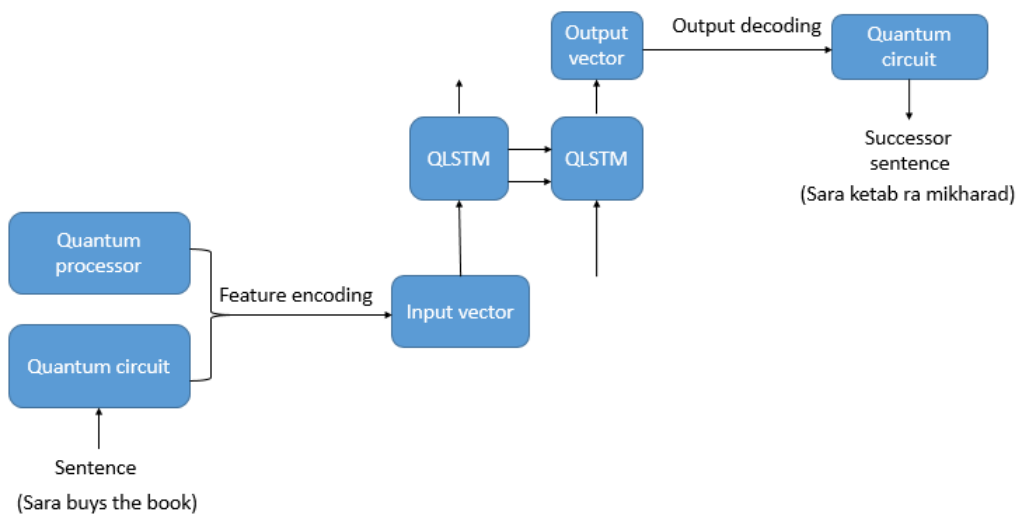


FIGURE 14. The proposed algorithm for translation of an English sentence to another language (e.g. Persian) via QLSTM.

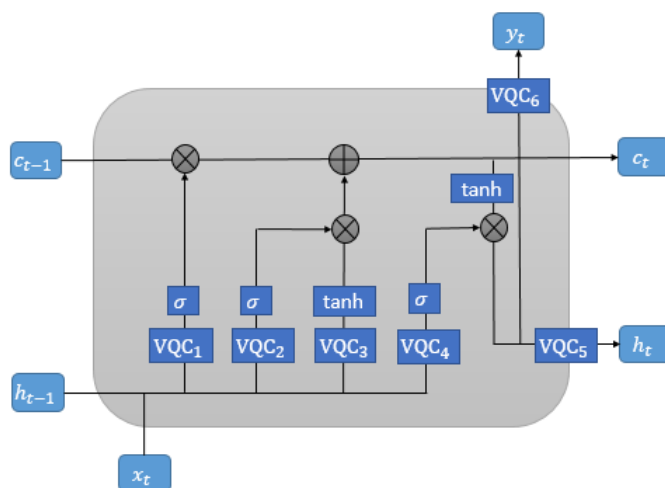


FIGURE 15. The structure of a QLSTM cell.

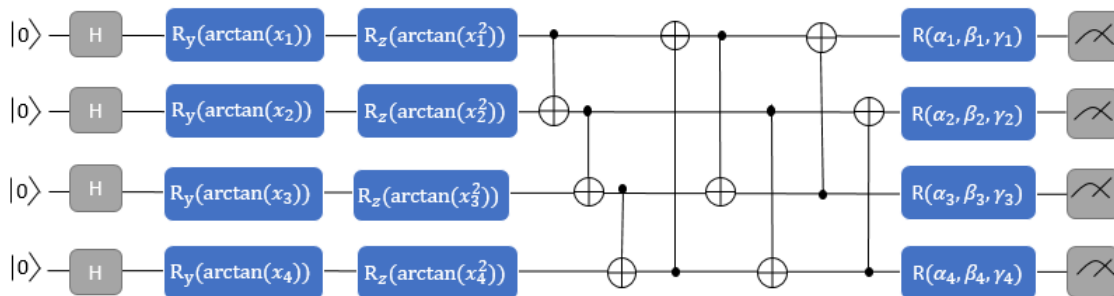


FIGURE 16. The architecture of VQC.

that the structures of these sentences are similar to English with similar circuits. Moreover, one may consider different parameters for synonymous sentences in different languages.

The method used to convert the DisCoCat diagrams into quantum circuits for our particular and simple examples, can be generalized to a recipe. Pregroups have been used to

TABLE 2. Examples of sentences in English and Persian.

English	فارسی
Sara buys the book.	سارا کتاب را می خرد.
Alice reads the book.	آلیس کتاب را می خواند.
Bob eats the cake.	باب کیک را می خورد.
The food my mother cooks, is delicious.	غذایی که مادرم می پزد، لذیذ است.
My mother cooks delicious food.	مادرم غذای لذیذ می پزد.

TABLE 3. Examples of sentences in two different languages.

Український (Ukrainian)	Spanish
Сара купує книгу (Sara kupuye knyhu)	Sara compra el libro
Аліса читає книгу (Alisa chytaye knyhu)	Alice lee el libro
Боб їсть торт (Bob yist' tort)	Bob se come el pastel

analyze the syntax of different languages, from English and French to Polish, and many more. For more references see [37], [38]. The proposed algorithm can be used for the translation of an English sentence to another language. In general, the recipe can work for all languages in the world, and not only for English and Persian.

A. QUANTUM ADVANTAGE

It has been shown that under certain conditions quantum algorithms for compositional QNLP demonstrate quadratic speedup over classical methods [39]. For example, an immediate advantage for quantum implementations of the DisCoCat diagrams [25] is gained by storing meaning vectors in quantum systems. Each n-qubit system with 2^n degrees of freedom, indicating that N-dimensional classical vectors can be stored in log2N qubits. Consider a corpus whose word-meaning space is given by a basis of the 2000 most common words, in the worst case one obtains the dramatic improvements, e.g. for one transitive verb 8x10^9 classical bits are required while only 33 qubits are required in QNLP, and thus for 10k transitive verbs 8x10^13 classical bits are required but only 47 qubits for this aim. Regarding the complexity comparisons for different closest vector algorithms, there is a quadratic improvement in scaling with M training vectors with the dimension of the vectors N, the complexity in classical regime is O(NM) while in quantum is O(sqrt(NM)log(M)) [39]. In QLSTM protocol, it has been shown that for certain testing cases, QLSTM converges faster or reaches a better accuracy than its classical counterpart, that paves the way toward implementation for sequence modeling on NISQ devices [31].

VI. RESULTS AND FUTURE DIRECTIONS

First of all, we highlight what have done well in the study. Then we give explanations and direct the future works.

- Developing a compositional vector-based semantics of a transitive sentence for a non-English language within a categorical framework.

- Comparing quantum circuits of two synonymous sentences in English and Persian.
- A quantum algorithm for language translation via quantum long short-term memory.
- The feasibility of a generalized algorithm for language translation for other languages.

In this paper, we have suggested a protocol for translating simple sentences from English to Persian via quantum natural language processing. First, we extended the compact categorical semantics to analyse meanings of positive transitive sentences in Persian language. It is necessary to introduce linear maps to represent the meaning of negative transitive sentences and grammatically more complex sentences in Persian. The demonstration of DisCoCat diagrams and quantum circuits of complicated sentences in Persian or any non-English language remains to be studied in future works. Here, the two sentences ‘Sara ketab ra mikharad’ (in Persian) and ‘Sara buys the book’ (in English) are instantiated as parametrized quantum circuits. The meaning of the two sentences are the same but the appearance of the obtained quantum circuits are different. These circuits need to be compiled correctly, thus it is necessary to introduce a test measurement at the terminal of the circuits to give almost similar results for the meaning of the synonymous sentences in different languages. One may use the compiler t|ket> to this aim, and run the circuits on the IBMQ and analyze the results. Finally, we propose an algorithm of using quantum long short-term memory for quantum natural language processing in order to translate the sentence above from English to Persian.

Here, we have proposed an algorithm in which one should first convert sentences to circuits and then to vectors. As a limitation, it does not convert the words to circuits and therefore not applicable for words, but it is good for texts which are composed of sentences. On the other hand, for implementation we should propose an algorithm that offer improvements to the quality of results, particularly for more complex sentences in terms of memory and computational requirements. As a future research prospect, one can perform different tests on IBMQ. For example, the meaning

classification task of sentences in Persian and performing the translation algorithm. Future directions include implementation of Q-NLP tasks such as sentence similarity of any non-English language and using real-world data with a pregroup parser. Bidirectional encoder-decoder model, which predicts sentences in English from the previous and subsequent texts, is another interesting future research direction. Moreover, one can investigate the uncertainty of the proposed model based on what is widely discussed in the literature [13], [40], [41].

## AUTHOR CONTRIBUTIONS

MA, VS and XZ have proposed the idea and consulted it with SSM and MZM, and all authors contributed to the development and completion of the idea, MA did the main part of the calculations, and the all authors participated in discussions and structure of the paper, and writing the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## REFERENCES

- [1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Cambridge, MA, USA: PEGA Publishing, 1995.
- [2] W. J. Hutchins, *Early Years in Machine Translation: Memoirs and Biographies of Pioneers*. Amsterdam, The Netherlands: John Benjamins Publishing, 2000.
- [3] R. P. Neco and M. L. Forcada, "Asynchronous translations with recurrent neural nets," in *Proc. IEEE Int. Conf. Neural Netw. (ICNN)*, vol. 4, Jun. 1997, pp. 2535–2540.
- [4] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, "A neural probabilistic language model," *J. Mach. Learn. Res.*, vol. 3, pp. 1137–1155, Feb. 2003.
- [5] N. Kalchbrenner and P. Blunsom, "Recurrent continuous translation models," *J. Mach. Learn. Res.*, vol. 3, no. 39, p. 413, 2013.
- [6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*. Red Hook, NY, USA: Curran Associates, 2017.
- [7] M. Popel, M. Tomkova, J. Tomek, L. Kaiser, J. Uszkoreit, O. Bojar, and Z. Žabokrtský, "Transforming machine translation: A deep learning system reaches news translation quality comparable to human professionals," *Nature Commun.*, vol. 11, no. 1, pp. 1–15, Dec. 2020, doi: 10.1038/s41467-020-18073-9.
- [8] M. Abdar, M. E. Basiri, J. Yin, M. Habibnezhad, G. Chi, S. Nemati, and S. Asadi, "Energy choices in Alaska: Mining people's perception and attitudes from geotagged tweets," *Renew. Sustain. Energy Rev.*, vol. 124, May 2020, Art. no. 109781.
- [9] M. E. Basiri, M. Abdar, A. Kabiri, S. Nemati, X. Zhou, F. Allahbakhshi, and N. Y. Yen, "Improving sentiment polarity detection through target identification," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 1, pp. 113–128, Feb. 2020.
- [10] M. E. Basiri, M. Abdar, M. A. Cifci, S. Nemati, and U. R. Acharya, "A novel method for sentiment classification of drug reviews using fusion of deep and machine learning techniques," *Knowl.-Based Syst.*, vol. 198, Jun. 2020, Art. no. 105949.
- [11] M. E. Basiri, A. Kabiri, M. Abdar, W. K. Mashwani, N. Y. Yen, and J. C. Hung, "The effect of aggregation methods on sentiment classification in Persian reviews," *Enterprise Inf. Syst.*, vol. 14, nos. 9–10, pp. 1394–1421, Nov. 2020.
- [12] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis," *Future Gener. Comput. Syst.*, vol. 115, pp. 279–294, Feb. 2021.
- [13] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Inf. Fusion*, vol. 76, pp. 243–297, Dec. 2021.
- [14] A. M. Turing, "Computing machinery and intelligence," *Mind*, vol. 59, no. 236, pp. 433–460, 1950.
- [15] T. B. Brown et al., "Language models are few-shot learners," 2020, *arXiv:2005.14165*. [Online]. Available: <https://arxiv.org/abs/2005.14165>
- [16] A. A. A. El-Latif, B. Abd-El-Atty, and M. Talha, "Robust encryption of quantum medical images," *IEEE Access*, vol. 6, pp. 1073–1081, 2017.
- [17] A. A. A. El-Latif, B. Abd-El-Atty, M. S. Hossain, M. A. Rahman, A. Alamri, and B. B. Gupta, "Efficient quantum information hiding for remote medical image sharing," *IEEE Access*, vol. 6, pp. 21075–21083, 2018.
- [18] M. Caleffi, "Optimal routing for quantum networks," *IEEE Access*, vol. 5, pp. 22299–22312, 2017.
- [19] Z. Qu, Z. Li, G. Xu, S. Wu, and X. Wang, "Quantum image steganography protocol based on quantum image expansion and Grover search algorithm," *IEEE Access*, vol. 7, pp. 50849–50857, 2019.
- [20] X.-B. Chen, Y.-L. Wang, G. Xu, and Y.-X. Yang, "Quantum network communication with a novel discrete-time quantum walk," *IEEE Access*, vol. 7, pp. 13634–13642, 2019.
- [21] W.-L. Chang, J.-C. Chen, W.-Y. Chung, C.-Y. Hsiao, R. Wong, and A. V. Vasilakos, "Quantum speedup and mathematical solutions of implementing bio-molecular solutions for the independent set problem on IBM quantum computers," *IEEE Trans. Nanobiosci.*, vol. 20, no. 3, pp. 354–376, Jul. 2021.
- [22] K. Meichanetzidis, S. Gogioso, G. de Felice, N. Chiappori, A. Toumi, and B. Coecke, "Quantum natural language processing on near-term quantum computers," 2020, *arXiv:2005.04147*. [Online]. Available: <http://arxiv.org/abs/2005.04147>
- [23] L. J. O'Riordan, M. Doyle, F. Baruffa, and V. Kannan, "A hybrid classical-quantum workflow for natural language processing," *Mach. Learn., Sci. Technol.*, vol. 2, no. 1, Dec. 2020, Art. no. 015011.
- [24] S. Abramsky and B. Coecke, "A categorical semantics of quantum protocols," in *Proc. 19th Annu. IEEE Symp. Log. Comput. Sci.*, Jul. 2004, pp. 415–425.
- [25] B. Coecke, M. Sadrzadeh, and S. Clark, "Mathematical foundations for a compositional distributional modal of meaning," *Linguistic Anal.*, vol. 36, pp. 345–384, Mar. 2010.
- [26] B. Coecke, "Quantum pictorialism," *Contemp. Phys.*, vol. 51, no. 1, pp. 59–83, 2010.
- [27] M. Sadrzadeh, "Pregroup analysis of Persian sentences," in *Computational Algebraic Approaches to Natural Language*, C. Casadio and J. Lambek, Eds. Monza, Italy: Polimetrica, 2006.
- [28] M. Sadrzadeh, S. Clark, and B. Coecke, "The Frobenius anatomy of word meanings I: Subject and object relative pronouns," *J. Log. Comput.*, vol. 23, no. 6, pp. 1293–1317, Dec. 2013.
- [29] B. Tyrrell, "Applying distributional compositional categorical models of meaning to language translation," in *Proc. Workshop Computational Approaches Phys., NLP, Social Sci. (CAPNS EPTCS)*, vol. 283, 2018, pp. 28–49.
- [30] B. Coecke, G. de Felice, K. Meichanetzidis, and A. Toumi, "Foundations for near-term quantum natural language processing," 2020, *arXiv:2012.03755*. [Online]. Available: <http://arxiv.org/abs/2012.03755>
- [31] S. Yen-Chi Chen, S. Yoo, and Y.-L. L. Fang, "Quantum long short-term memory," 2020, *arXiv:2009.01783*. [Online]. Available: <http://arxiv.org/abs/2009.01783>
- [32] B. Coecke and A. Kissinger, *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
- [33] E. Grefenstette and M. Sadrzadeh, "Experimental support for a categorical compositional distributional model of meaning," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2011, pp. 1394–1404.
- [34] B. Coecke and R. Duncan, "Interacting quantum observables: Categorical algebra and diagrammatics," *New J. Phys.*, vol. 13, no. 4, pp. 1394–1404, 2011.
- [35] R. Lorenz, A. Pearson, K. Meichanetzidis, D. Kartsaklis, and B. Coecke, "QNLP in practice: Running compositional models of meaning on a quantum computer," 2021, *arXiv:2102.12846*. [Online]. Available: <http://arxiv.org/abs/2102.12846>
- [36] E. Grefenstette, M. Sadrzadeh, S. Clark, B. Coecke, and S. Pulman, "Concrete sentence spaces for compositional distributional models of meaning," in *Proc. 9th Int. Conf. Comput. Semantics (IWCS)*, 2011, pp. 71–86.
- [37] J. Lambek, *From Word to Sentence*. Monza, Italy: Polimetrica, 2008.
- [38] J. Lambek and C. Casadio, Eds., *Computational Algebraic Approaches to Natural Language*. Monza, Italy: Polimetrica, 2006.



- [39] W. Zeng and B. Coecke, "Quantum algorithms for compositional natural language processing," *Electron. Proc. Theor. Comput. Sci.*, vol. 221, pp. 67–75, Aug. 2016.
- [40] M. Abdar, M. Samami, S. Dehghani Mahmoodabad, T. Doan, B. Mazouze, R. Hashemifesharaki, L. Liu, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, "Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning," *Comput. Biol. Med.*, vol. 135, Aug. 2021, Art. no. 104418.
- [41] M. Abdar, S. Salari, S. Qahremani, H.-K. Lam, F. Karray, S. Hussain, A. Khosravi, U. R. Acharya, and S. Nahavandi, "UncertaintyFuseNet: Robust uncertainty-aware hierarchical feature fusion with ensemble Monte Carlo dropout for COVID-19 detection," 2021, *arXiv:2105.08590*. [Online]. Available: <http://arxiv.org/abs/2105.08590>



**MINA ABBASZADE** received the Ph.D. degree in pure mathematics from Shahid Bahonar University of Kerman (SBUK), in 2018. She was a Postdoctoral Researcher at SBUK, from 2019 to 2020. She has been a Postdoctoral Researcher at Isfahan University of Technology, since 2020. Her research interests include category theory, quantum machine learning, and quantum neural networks.



**VAHID SALARI** was an Assistant Professor of physics at Isfahan University of Technology, from 2012 to 2019, and a Resident Researcher at the Institute for Research in Fundamental Sciences (IPM), Iran, from 2014 to 2017. He has been a Senior Researcher with the Department of Physical Chemistry, University of the Basque Country (UPV/EHU), Bilbao, Spain, since 2019. He is currently a Research Fellow at the Basque Center for Applied Mathematics (BCAM) and a Research

Associate at the Quantum Biology Laboratory, Howard University, Washington, DC, USA. His research interests include quantum technologies, brain studies, and quantum machine learning.



**SEYED SHAHIN MOUSAVI** is currently an Associate Professor of mathematics at Shahid Bahonar University of Kerman, Iran. His research interests include category theory, homological algebra, hyper structure theory, and picturing quantum processes.



**MARIAM ZOMORODI** received the B.Sc. degree in computer engineering from Iran University of Science and Technology, in 2002, the M.Sc. degree in computer engineering from Amirkabir University of Technology (Tehran Polytechnic), in 2007, and the Ph.D. degree in computer engineering from Shahid Beheshti University, in 2014. She is currently with the Department of Information and Communications Technology, Cracow University of Technology, and Ferdowsi University of Mashhad. Her current research interests include quantum computing, evolutionary computation, and machine learning.



**XUJUAN ZHOU** (Member, IEEE) received the Ph.D. degree from Queensland University of Technology, Australia, in 2009. She is currently a Senior Lecturer with the University of Southern Queensland, Australia. She has published articles in prestigious journals, such as *Pattern Recognition Letter*, *Decision Support Systems*, *Artificial Intelligence Review*, *Journal of Clinical Epidemiology*, and *Journal of Medical Internet Research*. Her research interests include machine learning, deep learning, health informatics, information retrieval/filtering, recommender systems, sentiment analysis, and natural language processing. She is an Editorial Board Member of *Web Intelligence* and a reviewer of many journals and conferences related to her research fields.

• • •