

Sequence-to-Action: End-to-End Semantic Graph Generation for Semantic Parsing

Bo Chen^{†‡}, Le Sun[†], Xianpei Han[†]

[†]State Key Laboratory of Computer Science
Institute of Software, Chinese Academy of Sciences, Beijing, China

[‡]University of Chinese Academy of Sciences, Beijing, China
{chenbo, sunle, xianpei}@iscas.ac.cn

Abstract

This paper proposes a neural semantic parsing approach – Sequence-to-Action, which models semantic parsing as an end-to-end semantic graph generation process. Our method simultaneously leverages the advantages from two recent promising directions of semantic parsing. Firstly, our model uses a semantic graph to represent the meaning of a sentence, which has a tight-coupling with knowledge bases. Secondly, by leveraging the powerful representation learning and prediction ability of neural network models, we propose a RNN model which can effectively map sentences to action sequences for semantic graph generation. Experiments show that our method achieves state-of-the-art performance on OVERNIGHT dataset and gets competitive performance on GEO and ATIS datasets.

1 Introduction

Semantic parsing aims to map natural language sentences to logical forms (Zelle and Mooney, 1996; Zettlemoyer and Collins, 2005; Wong and Mooney, 2007; Lu et al., 2008; Kwiatkowski et al., 2013). For example, the sentence “Which states border Texas?” will be mapped to answer (A, (state (A), next_to (A, stateid (texas))))).

A semantic parser needs two functions, one for structure prediction and the other for semantic grounding. Traditional semantic parsers are usually based on compositional grammar, such as CCG (Zettlemoyer and Collins, 2005, 2007), DCS (Liang et al., 2011), etc. These parsers compose structure using manually designed grammars, use lexicons for semantic grounding, and exploit fea-

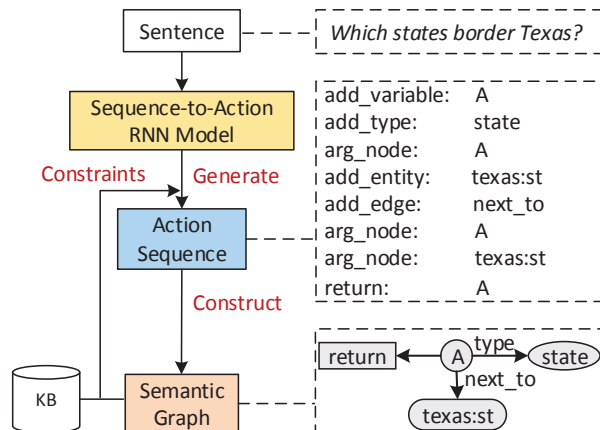


Figure 1: Overview of our method, with a demonstration example.

tures for candidate logical forms ranking. Unfortunately, it is challenging to design grammars and learn accurate lexicons, especially in wide-open domains. Moreover, it is often hard to design effective features, and its learning process is not end-to-end. To resolve the above problems, two promising lines of work have been proposed: Semantic graph-based methods and Seq2Seq methods.

Semantic graph-based methods (Reddy et al., 2014, 2016; Bast and Haussmann, 2015; Yih et al., 2015) represent the meaning of a sentence as a semantic graph (i.e., a sub-graph of a knowledge base, see example in Figure 1) and treat semantic parsing as a semantic graph matching/generation process. Compared with logical forms, semantic graphs have a tight-coupling with knowledge bases (Yih et al., 2015), and share many commonalities with syntactic structures (Reddy et al., 2014). Therefore both the structure and semantic constraints from knowledge bases can be easily exploited during parsing (Yih et al., 2015). The main challenge of semantic graph-based parsing is how to effectively construct the semantic graph of a sentence. Currently, semantic graphs

are either constructed by matching with patterns (Bast and Haussmann, 2015), transforming from dependency tree (Reddy et al., 2014, 2016), or via a staged heuristic search algorithm (Yih et al., 2015). These methods are all based on manually-designed, heuristic construction processes, making them hard to handle open/complex situations.

In recent years, RNN models have achieved success in sequence-to-sequence problems due to its strong ability on both representation learning and prediction, e.g., in machine translation (Cho et al., 2014). A lot of Seq2Seq models have also been employed for semantic parsing (Xiao et al., 2016; Dong and Lapata, 2016; Jia and Liang, 2016), where a sentence is parsed by translating it to linearized logical form using RNN models. There is no need for high-quality lexicons, manually-built grammars, and hand-crafted features. These models are trained end-to-end, and can leverage attention mechanism (Bahdanau et al., 2014; Luong et al., 2015) to learn soft alignments between sentences and logical forms.

In this paper, we propose a new neural semantic parsing framework – Sequence-to-Action, which can simultaneously leverage the advantages of semantic graph representation and the strong prediction ability of Seq2Seq models. Specifically, we model semantic parsing as an end-to-end semantic graph generation process. For example in Figure 1, our model will parse the sentence “Which states border Texas” by generating a sequence of actions [add_variable:A, add_type:state, ...]. To achieve the above goal, we first design an action set which can encode the generation process of semantic graph (including node actions such as add_variable, add_entity, add_type, edge actions such as add_edge, and operation actions such as argmin, argmax, count, sum, etc.). And then we design a RNN model which can generate the action sequence for constructing the semantic graph of a sentence. Finally we further enhance parsing by incorporating both structure and semantic constraints during decoding.

Compared with the manually-designed, heuristic generation algorithms used in traditional semantic graph-based methods, our sequence-to-action method generates semantic graphs using a RNN model, which is learned end-to-end from training data. Such a learnable, end-to-end generation makes our approach more effective and can fit to different situations.

Compared with the previous Seq2Seq semantic parsing methods, our sequence-to-action model predicts a sequence of semantic graph generation actions, rather than linearized logical forms. We find that the action sequence encoding can better capture structure and semantic information, and is more compact. And the parsing can be enhanced by exploiting structure and semantic constraints. For example, in GEO dataset, the action `add_edge:next_to` must subject to the semantic constraint that its arguments must be of type `state` and `state`, and the structure constraint that the edge `next_to` must connect two nodes to form a valid graph.

We evaluate our approach on three standard datasets: GEO (Zelle and Mooney, 1996), ATIS (He and Young, 2005) and OVERNIGHT (Wang et al., 2015b). The results show that our method achieves state-of-the-art performance on OVERNIGHT dataset and gets competitive performance on GEO and ATIS datasets.

The main contributions of this paper are summarized as follows:

- We propose a new semantic parsing framework – Sequence-to-Action, which models semantic parsing as an end-to-end semantic graph generation process. This new framework can synthesize the advantages of semantic graph representation and the prediction ability of Seq2Seq models.
- We design a sequence-to-action model, including an action set encoding for semantic graph generation and a Seq2Seq RNN model for action sequence prediction. We further enhance the parsing by exploiting structure and semantic constraints during decoding. Experiments validate the effectiveness of our method.

2 Sequence-to-Action Model for End-to-End Semantic Graph Generation

Given a sentence $X = x_1, \dots, x_{|X|}$, our sequence-to-action model generates a sequence of actions $Y = y_1, \dots, y_{|Y|}$ for constructing the correct semantic graph. Figure 2 shows an example. The conditional probability $P(Y|X)$ used in our

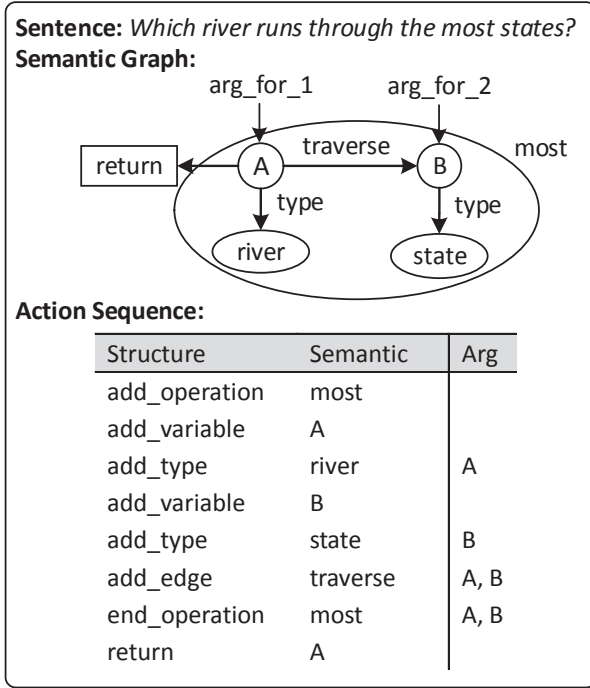


Figure 2: An example of a sentence paired with its semantic graph, together with the action sequence for semantic graph generation.

model is decomposed as follows:

$$P(Y|X) = \prod_{t=1}^{|Y|} P(y_t|y_{<t}, X) \quad (1)$$

where $y_{<t} = y_1, \dots, y_{t-1}$.

To achieve the above goal, we need: 1) an action set which can encode semantic graph generation process; 2) an encoder which encodes natural language input X into a vector representation, and a decoder which generates $y_1, \dots, y_{|Y|}$ conditioned on the encoding vector. In following we describe them in detail.

2.1 Actions for Semantic Graph Generation

Generally, a semantic graph consists of nodes (including variables, entities, types) and edges (semantic relations), with some universal operations (e.g., *argmax*, *argmin*, *count*, *sum*, and *not*). To generate a semantic graph, we define six types of actions as follows:

Add Variable Node: This kind of actions denotes adding a variable node to semantic graph. In most cases a variable node is a return node (e.g., *which*, *what*), but can also be an intermediate variable node. We represent this kind of action as `add_variable:A`, where A is the identifier of the variable node.

Add Entity Node: This kind of actions denotes adding an entity node (e.g., *Texas*, *New York*) and is represented as `add_entity_node:texas`. An entity node corresponds to an entity in knowledge bases.

Add Type Node: This kind of actions denotes adding a type node (e.g., *state*, *city*). We represent them as `add_type_node:state`.

Add Edge: This kind of actions denotes adding an edge between two nodes. An edge is a binary relation in knowledge bases. This kind of actions is represented as `add_edge:next_to`.

Operation Action: This kind of actions denotes adding an operation. An operation can be *argmax*, *argmin*, *count*, *sum*, *not*, et al. Because each operation has a scope, we define two actions for an operation, one is operation start action, represented as `start_operation:most`, and the other is operation end action, represented as `end_operation:most`. The subgraph within the start and end operation actions is its scope.

Argument Action: Some above actions need argument information. For example, which nodes the `add_edge:next_to` action should connect to. In this paper, we design argument actions for `add_type`, `add_edge` and operation actions, and the argument actions should be put directly after its main action.

For `add_type` actions, we put an argument action to indicate which node this type node should constrain. The argument can be a variable node or an entity node. An argument action for a type node is represented as `arg:A`.

For `add_edge` action, we use two argument actions: `arg1_node` and `arg2_node`, and they are represented as `arg1_node:A` and `arg2_node:B`.

We design argument actions for different operations. For `operation:sum`, there are three arguments: `arg-for`, `arg-in` and `arg-return`. For `operation:count`, they are `arg-for` and `arg-return`. There are two `arg-for` arguments for `operation:most`.

We can see that each action encodes both structure and semantic information, which makes it easy to capture more information for parsing and can be tightly coupled with knowledge base. Furthermore, we find that action sequence encoding is more compact than linearized logical form (See Section 4.4 for more details).

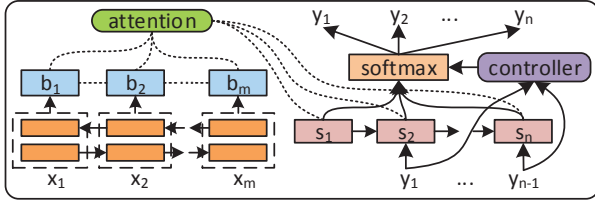


Figure 3: Our attention-based Sequence-to-Action RNN model, with a controller for incorporating constraints.

2.2 Neural Sequence-to-Action Model

Based on the above action encoding mechanism, this section describes our encoder-decoder model for mapping sentence to action sequence. Specifically, similar to the RNN model in [Jia and Liang \(2016\)](#), this paper employs the attention-based sequence-to-sequence RNN model. Figure 3 presents the overall structure.

Encoder: The encoder converts the input sequence x_1, \dots, x_m to a sequence of context-sensitive vectors b_1, \dots, b_m using a bidirectional RNN ([Bahdanau et al., 2014](#)). Firstly each word x_i is mapped to its embedding vector, then these vectors are fed into a forward RNN and a backward RNN. The sequence of hidden states h_1, \dots, h_m are generated by recurrently applying the recurrence:

$$h_i = LSTM(\phi^{(x)}(x_i), h_{i-1}). \quad (2)$$

The recurrence takes the form of LSTM ([Hochreiter and Schmidhuber, 1997](#)). Finally, for each input position i , we define its context-sensitive embedding as $b_i = [h_i^F, h_i^B]$.

Decoder: This paper uses the classical attention-based decoder ([Bahdanau et al., 2014](#)), which generates action sequence y_1, \dots, y_n , one action at a time. At each time step j , it writes y_j based on the current hidden state s_j , then updates the hidden state to s_{j+1} based on s_j and y_j . The decoder is formally defined by the following equations:

$$s_1 = \tanh(W^{(s)}[h_m^F, h_1^B]) \quad (3)$$

$$e_{ji} = s_j^T W^{(a)} b_i \quad (4)$$

$$a_{ji} = \frac{\exp(e_{ji})}{\sum_{i'=1}^m \exp(e_{ji'})} \quad (5)$$

$$c_j = \sum_{i=1}^m a_{ji} b_i \quad (6)$$

$$P(y_j = w | x, y_{1:j-1}) \propto \exp(U_w[s_j, c_j]) \quad (7)$$

$$s_{j+1} = LSTM([\phi^{(y)}(y_j), c_j], s_j) \quad (8)$$

where the normalized attention scores a_{ji} defines the probability distribution over input words, indicating the attention probability on input word i at time j ; e_{ji} is un-normalized attention score. To incorporate constraints during decoding, an extra controller component is added and its details will be described in Section 3.3.

Action Embedding. The above decoder needs the embedding of each action. As described above, each action has two parts, one for structure (e.g., `add.edge`), and the other for semantic (e.g., `next.to`). As a result, actions may share the same structure or semantic part, e.g., `add.edge:next.to` and `add.edge:loc` have the same structure part, and `add.node:A` and `arg.node:A` have the same semantic part. To make parameters more compact, we first embed the structure part and the semantic part independently, then concatenate them to get the final embedding. For instance, $\phi^{(y)}(\text{add.edge:next.to}) = [\phi_{struct}^{(y)}(\text{add.edge}), \phi_{sem}^{(y)}(\text{next.to})]$. The action embeddings $\phi^{(y)}$ are learned during training.

3 Constrained Semantic Parsing using Sequence-to-Action Model

In this section, we describe how to build a neural semantic parser using sequence-to-action model. We first describe the training and the inference of our model, and then introduce how to incorporate structure and semantic constraints during decoding.

3.1 Training

Parameter Estimation. The parameters of our model include RNN parameters $W^{(s)}$, $W^{(a)}$, U_w , word embeddings $\phi^{(x)}$, and action embeddings $\phi^{(y)}$. We estimate these parameters from training data. Given a training example with a sentence X and its action sequence Y , we maximize the likelihood of the generated sequence of actions given X . The objective function is:

$$\sum_{i=1}^n \log P(Y_i | X_i) \quad (9)$$

Standard stochastic gradient descent algorithm is employed to update parameters.

Logical Form to Action Sequence. Currently, most datasets of semantic parsing are labeled with logical forms. In order to train our model, we

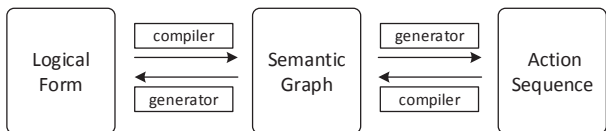


Figure 4: The procedure of converting between logical form and action sequence.

convert logical forms to action sequences using semantic graph as an intermediate representation (See Figure 4 for an overview). Concretely, we transform logical forms into semantic graphs using a depth-first-search algorithm from root, and then generate the action sequence using the same order. Specifically, entities, variables and types are nodes; relations are edges. Conversely we can convert action sequence to logical form similarly. Based on the above algorithm, action sequences can be transformed into logical forms in a deterministic way, and the same for logical forms to action sequences.

Mechanisms for Handling Entities. Entities play an important role in semantic parsing (Yih et al., 2015). In Dong and Lapata (2016), entities are replaced with their types and unique IDs. In Jia and Liang (2016), entities are generated via attention-based copying mechanism helped with a lexicon. This paper implements both mechanisms and compares them in experiments.

3.2 Inference

Given a new sentence X , we predict action sequence by:

$$Y^* = \operatorname{argmax}_Y P(Y|X) \quad (10)$$

where Y represents action sequence, and $P(Y|X)$ is computed using Formula (1). Beam search is used for best action sequence decoding. Semantic graph and logical form can be derived from Y^* as described in above.

3.3 Incorporating Constraints in Decoding

For decoding, we generate action sequentially. It is obviously that the next action has a strong correlation with the partial semantic graph generated to current, and illegal actions can be filtered using structure and semantic constraints. Specifically, we incorporate constraints in decoding using a controller. This procedure has two steps: 1) the controller constructs partial semantic graph using the actions generated to current; 2) the controller checks whether a new generated action can meet

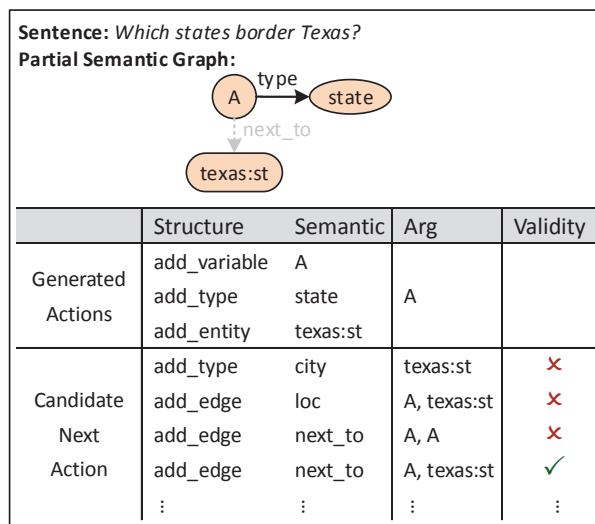


Figure 5: A demonstration of illegal action filtering using constraints. The graph in color is the constructed semantic graph to current.

all structure/semantic constraints using the partial semantic graph.

Structure Constraints. The structure constraints ensure action sequence will form a connected acyclic graph. For example, there must be two argument nodes for an edge, and the two argument nodes should be different (The third candidate next action in Figure 5 violates this constraint). This kind of constraints are domain-independent. The controller encodes structure constraints as a set of rules.

Semantic Constraints. The semantic constraints ensure the constructed graph must follow the schema of knowledge bases. Specifically, we model two types of semantic constraints. One is selectional preference constraints where the argument types of a relation should follow knowledge base schemas. For example, in GEO dataset, relation *next_to*'s *arg1* and *arg2* should both be a *state*. The second is type conflict constraints, i.e., an entity/variable node's type must be consistent, i.e., a node cannot be both of type *city* and *state*. Semantic constraints are domain-specific and are automatically extracted from knowledge base schemas. The controller encodes semantic constraints as a set of rules.

4 Experiments

In this section, we assess the performance of our method and compare it with previous methods.

4.1 Datasets

We conduct experiments on three standard datasets: GEO, ATIS and OVERNIGHT.

GEO contains natural language questions about US geography paired with corresponding Prolog database queries. Following Zettlemoyer and Collins (2005), we use the standard 600/280 instance splits for training/test.

ATIS contains natural language questions of a flight database, with each question is annotated with a lambda calculus query. Following Zettlemoyer and Collins (2007), we use the standard 4473/448 instance splits for training/test.

OVERNIGHT contains natural language paraphrases paired with logical forms across eight domains. We evaluate on the standard train/test splits as Wang et al. (2015b).

4.2 Experimental Settings

Following the experimental setup of Jia and Liang (2016): we use 200 hidden units and 100-dimensional word vectors for sentence encoding. The dimensions of action embedding are tuned on validation datasets for each corpus. We initialize all parameters by uniformly sampling within the interval [-0.1, 0.1]. We train our model for a total of 30 epochs with an initial learning rate of 0.1, and halve the learning rate every 5 epochs after epoch 15. We replace word vectors for words occurring only once with an universal word vector. The beam size is set as 5. Our model is implemented in Theano (Bergstra et al., 2010), and the codes and settings are released on Github: <https://github.com/dongpobeyond/Seq2Act>.

We evaluate different systems using the standard accuracy metric, and the accuracies on different datasets are obtained as same as Jia and Liang (2016).

4.3 Overall Results

We compare our method with state-of-the-art systems on all three datasets. Because all systems using the same training/test splits, we directly use the reported best performances from their original papers for fair comparison.

For our method, we train our model with three settings: the first one is the basic sequence-to-action model without constraints – Seq2Act; the second one adds structure constraints in decoding – Seq2Act (+C1); the third one is the full model which adds both structure and semantic

	GEO	ATIS
Previous Work		
Zettlemoyer and Collins (2005)	79.3	–
Zettlemoyer and Collins (2007)	86.1	84.6
Kwiatkowski et al. (2010)	88.9	–
Kwiatkowski et al. (2011)	88.6	82.8
Liang et al. (2011)* (+lexicon)	91.1	–
Poon (2013)	–	83.5
Zhao et al. (2015)	88.9	84.2
Rabinovich et al. (2017)	87.1	85.9
Seq2Seq Models		
Jia and Liang (2016)	85.0	76.3
Jia and Liang (2016)* (+data)	89.3	83.3
Dong and Lapata (2016): 2Seq	84.6	84.2
Dong and Lapata (2016): 2Tree	87.1	84.6
Our Models		
Seq2Act	87.5	84.6
Seq2Act (+C1)	88.2	85.0
Seq2Act (+C1+C2)	88.9	85.5

Table 1: Test accuracies on GEO and ATIS datasets, where * indicates systems with extra-resources are used.

constraints – Seq2Act (+C1+C2). Semantic constraints (C2) are stricter than structure constraints (C1). Therefore we set that C1 should be first met for C2 to be met. So in our experiments we add constraints incrementally. The overall results are shown in Table 1-2. From the overall results, we can see that:

1) By synthesizing the advantages of semantic graph representation and the prediction ability of Seq2Seq model, our method achieves state-of-the-art performance on OVERNIGHT dataset, and gets competitive performance on GEO and ATIS dataset. In fact, on GEO our full model (Seq2Act+C1+C2) also gets the best test accuracy of 88.9 if under the same settings, which only falls behind Liang et al. (2011)* which uses extra hand-crafted lexicons and Jia and Liang (2016)* which uses extra augmented training data. On ATIS our full model gets the second best test accuracy of 85.5, which only falls behind Rabinovich et al. (2017) which uses a supervised attention strategy. On OVERNIGHT, our full model gets state-of-the-art accuracy of 79.0, which even outperforms Jia and Liang (2016)* with extra augmented training data.

2) Compared with the linearized logical form representation used in previous Seq2Seq baselines, our action sequence encoding is more effective for semantic parsing. On all three datasets,

	Soc.	Blo.	Bas.	Res.	Cal.	Hou.	Pub.	Rec.	Avg.
Previous Work									
Wang et al. (2015b)	48.2	41.9	46.3	75.9	74.4	54.0	59.0	70.8	58.8
Seq2Seq Models									
Xiao et al. (2016)	80.0	55.6	80.5	80.1	75.0	61.9	75.8	–	72.7
Jia and Liang (2016)	81.4	58.1	85.2	76.2	78.0	71.4	76.4	79.6	75.8
Jia and Liang (2016)* (+data)	79.6	60.2	87.5	79.5	81.0	72.5	78.3	81.0	77.5
Our Models									
Seq2Act	81.4	60.4	87.5	79.8	81.0	73.0	79.5	81.5	78.0
Seq2Act (+C1)	81.8	60.9	88.0	80.1	81.0	73.5	80.1	82.0	78.4
Seq2Act (+C1+C2)	82.1	61.4	88.2	80.7	81.5	74.1	80.7	82.9	79.0

Table 2: Test accuracies on OVERNIGHT dataset, which includes eight domains: Social, Blocks, Basketball, Restaurants, Calendar, Housing, Publications, and Recipes.

our basic Seq2Act model gets better results than all Seq2Seq baselines. On GEO, the Seq2Act model achieve test accuracy of 87.5, better than the best accuracy 87.1 of Seq2Seq baseline. On ATIS, the Seq2Act model obtains a test accuracy of 84.6, the same as the best Seq2Seq baseline. On OVERNIGHT, the Seq2Act model gets a test accuracy of 78.0, better than the best Seq2Seq baseline gets 77.5. We argue that this is because our action sequence encoding is more compact and can capture more information.

3) Structure constraints can enhance semantic parsing by ensuring the validity of graph using the generated action sequence. In all three datasets, Seq2Act (+C1) outperforms the basic Seq2Act model. This is because a part of illegal actions will be filtered during decoding.

4) By leveraging knowledge base schemas during decoding, semantic constraints are effective for semantic parsing. Compared to Seq2Act and Seq2Act (+C1), the Seq2Act (+C1+C2) gets the best performance on all three datasets. This is because semantic constraints can further filter semantic illegal actions using selectional preference and consistency between types.

4.4 Detailed Analysis

Effect of Entity Handling Mechanisms. This paper implements two entity handling mechanisms – *Replacing* (Dong and Lapata, 2016) which identifies entities and then replaces them with their types and IDs, and attention-based *Copying* (Jia and Liang, 2016). To compare the above two mechanisms, we train and test with our full model and the results are shown in Table 3. We can see that, *Replacing* mechanism outperforms *Copying* in all three datasets. This is because *Replacing* is done

	<i>Replacing</i>	<i>Copying</i>
GEO	88.9	88.2
ATIS	85.5	84.0
OVERNIGHT	79.0	77.9

Table 3: Test accuracies of Seq2Act (+C1+C2) on GEO, ATIS, and OVERNIGHT of two entity handling mechanisms.

	Logical Form	Action Sequence
GEO	28.2	18.2
ATIS	28.4	25.8
OVERNIGHT	46.6	33.3

Table 4: Average length of logical forms and action sequences on three datasets. On OVERNIGHT, we average across all eight domains.

in preprocessing, while attention-based *Copying* is done during parsing and needs additional copy mechanism.

Linearized Logical Form vs. Action Sequence. Table 4 shows the average length of linearized logical forms used in previous Seq2Seq models and the action sequences of our model on all three datasets. As we can see, action sequence encoding is more compact than linearized logical form encoding: action sequence is shorter on all three datasets, 35.5%, 9.2% and 28.5% reduction in length respectively. The main advantage of a shorter/compact encoding is that it will reduce the influence of long distance dependency problem.

4.5 Error Analysis

We perform error analysis on results and find there are mainly two types of errors.

Unseen/Informal Sentence Structure. Some test sentences have unseen syntactic structures. For example, the first case in Table 5 has an unseen

Error Types	Examples
Un-covered Sentence Structure	Sentence: <i>Iowa borders how many states?</i> (Formal Form: <i>How many states does Iowa border?</i>) Gold Parse: <code>answer(A, count(B, (const(C, stateid(iowa)), next_to(C, B), state(B)), A))</code> Predicted Parse: <code>answer(A, count(B, state(B), A))</code>
Under-Mapping	Sentence: <i>Please show me first class flights from indianapolis to memphis one way leaving before 10am</i> Gold Parse: <code>(lambda x (and (flight x) (oneway x) (class_type x first:cl) (< (departure_time x) 1000:ti) (from x indianapolis:ci) (to x memphis:ci)))</code> Predicted Parse: <code>(lambda x (and (flight x) (oneway x) (< (departure_time x) 1000:ti) (from x indianapolis:ci) (to x memphis:ci)))</code>

Table 5: Some examples for error analysis. Each example includes the sentence for parsing, with gold parse and predicted parse from our model.

and informal structure, where entity word “*Iowa*” and relation word “*borders*” appear ahead of the question words “*how many*”. For this problem, we can employ sentence rewriting or paraphrasing techniques (Chen et al., 2016; Dong et al., 2017) to transform unseen sentence structures into normal ones.

Under-Mapping. As Dong and Lapata (2016) discussed, the attention model does not take the alignment history into consideration, makes some words are ignored during parsing. For example in the second case in Table 5, “*first class*” is ignored during the decoding process. This problem can be further solved using explicit word coverage models used in neural machine translation (Tu et al., 2016; Cohn et al., 2016)

5 Related Work

Semantic parsing has received significant attention for a long time (Kate and Mooney, 2006; Clarke et al., 2010; Krishnamurthy and Mitchell, 2012; Artzi and Zettlemoyer, 2013; Berant and Liang, 2014; Quirk et al., 2015; Artzi et al., 2015; Reddy et al., 2017). Traditional methods are mostly based on the principle of compositional semantics, which first trigger predicates using lexicons and then compose them using grammars. The prominent grammars include SCFG (Wong and Mooney, 2007; Li et al., 2015), CCG (Zettlemoyer and Collins, 2005; Kwiatkowski et al., 2011; Cai and Yates, 2013), DCS (Liang et al., 2011; Berant et al., 2013), etc. As discussed above, the main drawback of grammar-based methods is that they rely on high-quality lexicons, manually-built grammars, and hand-crafted features.

In recent years, one promising direction of semantic parsing is to use semantic graph as representation. Thus semantic parsing is modeled as a semantic graph generation process. Ge and Mooney (2009) build semantic graph by trans-

forming syntactic tree. Bast and Haussmann (2015) identify the structure of a semantic query using three pre-defined patterns. Reddy et al. (2014, 2016) use Freebase-based semantic graph representation, and convert sentences to semantic graphs using CCG or dependency tree. Yih et al. (2015) generate semantic graphs using a staged heuristic search algorithm. These methods are all based on manually-designed, heuristic generation process, which may suffer from syntactic parse errors (Ge and Mooney, 2009; Reddy et al., 2014, 2016), structure mismatch (Chen et al., 2016), and are hard to deal with complex sentences (Yih et al., 2015).

One other direction is to employ neural Seq2Seq models, which models semantic parsing as an end-to-end, sentence to logical form machine translation problem. Dong and Lapata (2016), Jia and Liang (2016) and Xiao et al. (2016) transform word sequence to linearized logical forms. One main drawback of these methods is that it is hard to capture and exploit structure and semantic constraints using linearized logical forms. Dong and Lapata (2016) propose a Seq2Tree model to capture the hierarchical structure of logical forms.

It has been shown that structure and semantic constraints are effective for enhancing semantic parsing. Krishnamurthy et al. (2017) use type constraints to filter illegal tokens. Liang et al. (2017) adopt a Lisp interpreter with pre-defined functions to produce valid tokens. Iyyer et al. (2017) adopt type constraints to generate valid actions. Inspired by these approaches, we also incorporate both structure and semantic constraints in our neural sequence-to-action model.

Transition-based approaches are important in both dependency parsing (Nivre, 2008; Henderson et al., 2013) and AMR parsing (Wang et al., 2015a). In semantic parsing, our method has a tight-coupling with knowledge bases, and con-

straints can be exploited for more accurate decoding. We believe this can also be used to enhance previous transition based methods and may also be used in other parsing tasks, e.g., AMR parsing.

6 Conclusions

This paper proposes Sequence-to-Action, a method which models semantic parsing as an end-to-end semantic graph generation process. By leveraging the advantages of semantic graph representation and exploiting the representation learning and prediction ability of Seq2Seq models, our method achieved significant performance improvements on three datasets. Furthermore, structure and semantic constraints can be easily incorporated in decoding to enhance semantic parsing.

For future work, to solve the problem of the lack of training data, we want to design weakly supervised learning algorithm using denotations (QA pairs) as supervision. Furthermore, we want to collect labeled data by designing an interactive UI for annotation assist like (Yih et al., 2016), which uses semantic graphs to annotate the meaning of sentences, since semantic graph is more natural and can be easily annotated without the need of expert knowledge.

Acknowledgments

This research work is supported by the National Key Research and Development Program of China under Grant No.2017YFB1002104; and the National Natural Science Foundation of China under Grants no. 61572477 and 61772505. Moreover, we sincerely thank the reviewers for their valuable comments.

References

Yoav Artzi, Kenton Lee, and Luke Zettlemoyer. 2015. [Broad-coverage ccg semantic parsing with amr](#). In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Lisbon, Portugal, pages 1699–1710. <http://aclweb.org/anthology/D15-1198>.

Yoav Artzi and Luke Zettlemoyer. 2013. Weakly supervised learning of semantic parsers for mapping instructions to actions. *Transactions of the Association for Computational Linguistics* 1(1):49–62.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. [Neural machine translation by jointly learning to align and translate](#). *CoRR* abs/1409.0473. <http://arxiv.org/abs/1409.0473>.

Hannah Bast and Elmar Haussmann. 2015. [More accurate question answering on freebase](#). In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, CIKM 2015, Melbourne, VIC, Australia, October 19 - 23, 2015*. pages 1431–1440. <https://doi.org/10.1145/2806416.2806472>.

Jonathan Berant, Andrew Chou, Roy Frostig, and Percy Liang. 2013. [Semantic parsing on Freebase from question-answer pairs](#). In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Seattle, Washington, USA, pages 1533–1544. <http://www.aclweb.org/anthology/D13-1160>.

Jonathan Berant and Percy Liang. 2014. [Semantic parsing via paraphrasing](#). In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Baltimore, Maryland, pages 1415–1425. <http://www.aclweb.org/anthology/P14-1133>.

James Bergstra, Olivier Breuleux, Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume Desjardins, Joseph Turian, David Warde-Farley, and Yoshua Bengio. 2010. Theano: A cpu and gpu math compiler in python. In *Proc. 9th Python in Science Conf.* pages 1–7.

Qingqing Cai and Alexander Yates. 2013. [Large-scale semantic parsing via schema matching and lexicon extension](#). In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Sofia, Bulgaria, pages 423–433. <http://www.aclweb.org/anthology/P13-1042>.

Bo Chen, Le Sun, Xianpei Han, and Bo An. 2016. [Sentence rewriting for semantic parsing](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 766–777. <http://www.aclweb.org/anthology/P16-1073>.

Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. [Learning phrase representations using rnn encoder-decoder for statistical machine translation](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Doha, Qatar, pages 1724–1734. <http://www.aclweb.org/anthology/D14-1179>.

James Clarke, Dan Goldwasser, Ming-Wei Chang, and Dan Roth. 2010. [Driving semantic parsing from the world’s response](#). In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*. Association for Computational Linguistics, Uppsala, Sweden, pages 18–27. <http://www.aclweb.org/anthology/W10-2903>.

- Trevor Cohn, Cong Duy Vu Hoang, Ekaterina Vymolova, Kaisheng Yao, Chris Dyer, and Gholamreza Haffari. 2016. Incorporating structural alignment biases into an attentional neural translation model. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, San Diego, California, pages 876–885. <http://www.aclweb.org/anthology/N16-1102>.
- Li Dong and Mirella Lapata. 2016. Language to logical form with neural attention. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 33–43. <http://www.aclweb.org/anthology/P16-1004>.
- Li Dong, Jonathan Mallinson, Siva Reddy, and Mirella Lapata. 2017. Learning to paraphrase for question answering. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Copenhagen, Denmark, pages 875–886. <https://www.aclweb.org/anthology/D17-1091>.
- Ruifang Ge and Raymond Mooney. 2009. Learning a compositional semantic parser using an existing syntactic parser. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*. Association for Computational Linguistics, Suntec, Singapore, pages 611–619. <http://www.aclweb.org/anthology/P/P09/P09-1069>.
- Yulan He and Steve Young. 2005. Semantic processing using the hidden vector state model. *Computer Speech Language* 19(1):85 – 106. <https://doi.org/https://doi.org/10.1016/j.csl.2004.03.001>.
- James Henderson, Paola Merlo, Ivan Titov, and Gabriele Musillo. 2013. Multilingual joint parsing of syntactic and semantic dependencies with a latent variable model. *Comput. Linguist.* 39(4):949–998. http://dx.doi.org/10.1162/COLL_a0158.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Comput.* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Mohit Iyyer, Wen-tau Yih, and Ming-Wei Chang. 2017. Search-based neural structured learning for sequential question answering. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Vancouver, Canada, pages 1821–1831. <http://aclweb.org/anthology/P17-1167>.
- Robin Jia and Percy Liang. 2016. Data recombination for neural semantic parsing. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 12–22. <http://www.aclweb.org/anthology/P16-1002>.
- Rohit J. Kate and Raymond J. Mooney. 2006. Using string-kernels for learning semantic parsers. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Sydney, Australia, pages 913–920. <https://doi.org/10.3115/1220175.1220290>.
- Jayant Krishnamurthy, Pradeep Dasigi, and Matt Gardner. 2017. Neural semantic parsing with type constraints for semi-structured tables. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Copenhagen, Denmark, pages 1516–1526. <https://www.aclweb.org/anthology/D17-1160>.
- Jayant Krishnamurthy and Tom Mitchell. 2012. Weakly supervised training of semantic parsers. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Association for Computational Linguistics, Jeju Island, Korea, pages 754–765. <http://www.aclweb.org/anthology/D12-1069>.
- Tom Kwiatkowski, Luke Zettlemoyer, Sharon Goldwater, and Mark Steedman. 2010. Inducing probabilistic CCG grammars from logical form with higher-order unification. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Cambridge, MA, pages 1223–1233. <http://www.aclweb.org/anthology/D10-1119>.
- Tom Kwiatkowski, Eunsol Choi, Yoav Artzi, and Luke Zettlemoyer. 2013. Scaling semantic parsers with on-the-fly ontology matching. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Seattle, Washington, USA, pages 1545–1556. <http://www.aclweb.org/anthology/D13-1161>.
- Tom Kwiatkowski, Luke Zettlemoyer, Sharon Goldwater, and Mark Steedman. 2011. Lexical generalization in ccg grammar induction for semantic parsing. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Edinburgh, Scotland, UK., pages 1512–1523. <http://www.aclweb.org/anthology/D11-1140>.
- Junhui Li, Muhua Zhu, Wei Lu, and Guodong Zhou. 2015. Improving semantic parsing with enriched synchronous context-free grammar. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Lisbon, Portugal, pages 1455–1465. <http://aclweb.org/anthology/D15-1170>.

- Chen Liang, Jonathan Berant, Quoc Le, Kenneth D. Forbus, and Ni Lao. 2017. *Neural symbolic machines: Learning semantic parsers on freebase with weak supervision*. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Vancouver, Canada, pages 23–33. <http://aclweb.org/anthology/P17-1003>.
- Percy Liang, Michael Jordan, and Dan Klein. 2011. *Learning dependency-based compositional semantics*. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Portland, Oregon, USA, pages 590–599. <http://www.aclweb.org/anthology/P11-1060>.
- Wei Lu, Hwee Tou Ng, Wee Sun Lee, and Luke S. Zettlemoyer. 2008. *A generative model for parsing natural language to meaning representations*. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Honolulu, Hawaii, pages 783–792. <http://www.aclweb.org/anthology/D08-1082>.
- Thang Luong, Hieu Pham, and Christopher D. Manning. 2015. *Effective approaches to attention-based neural machine translation*. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Lisbon, Portugal, pages 1412–1421. <http://aclweb.org/anthology/D15-1166>.
- Joakim Nivre. 2008. *Algorithms for deterministic incremental dependency parsing*. *Comput. Linguist.* 34(4):513–553. <http://dx.doi.org/10.1162/coli.07-056-R1-07-027>.
- Hoifung Poon. 2013. *Grounded unsupervised semantic parsing*. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Sofia, Bulgaria, pages 933–943. <http://www.aclweb.org/anthology/P13-1092>.
- Chris Quirk, Raymond Mooney, and Michel Galley. 2015. *Language to code: Learning semantic parsers for if-this-then-that recipes*. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Beijing, China, pages 878–888. <http://www.aclweb.org/anthology/P15-1085>.
- Maxim Rabinovich, Mitchell Stern, and Dan Klein. 2017. *Abstract syntax networks for code generation and semantic parsing*. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Vancouver, Canada, pages 1139–1149. <http://aclweb.org/anthology/P17-1105>.
- Siva Reddy, Mirella Lapata, and Mark Steedman. 2014. *Large-scale semantic parsing without question-answer pairs*. *Transactions of the Association for Computational Linguistics* 2:377–392. <http://aclweb.org/anthology/Q14-1030>.
- Siva Reddy, Oscar Täckström, Michael Collins, Tom Kwiatkowski, Dipanjan Das, Mark Steedman, and Mirella Lapata. 2016. *Transforming Dependency Structures to Logical Forms for Semantic Parsing*. *Transactions of the Association for Computational Linguistics* 4:127–140. <http://sivareddy.in/papers/reddy2016transforming.pdf>.
- Siva Reddy, Oscar Täckström, Slav Petrov, Mark Steedman, and Mirella Lapata. 2017. *Universal semantic parsing*. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Copenhagen, Denmark, pages 89–101. <https://www.aclweb.org/anthology/D17-1009>.
- Zhaopeng Tu, Zhengdong Lu, Yang Liu, Xiaohua Liu, and Hang Li. 2016. *Modeling coverage for neural machine translation*. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 76–85. <http://www.aclweb.org/anthology/P16-1008>.
- Chuan Wang, Nianwen Xue, and Sameer Pradhan. 2015a. *A transition-based algorithm for amr parsing*. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Denver, Colorado, pages 366–375. <http://www.aclweb.org/anthology/N15-1040>.
- Yushi Wang, Jonathan Berant, and Percy Liang. 2015b. *Building a semantic parser overnight*. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Beijing, China, pages 1332–1342. <http://www.aclweb.org/anthology/P15-1129>.
- Yuk Wah Wong and Raymond Mooney. 2007. *Learning synchronous grammars for semantic parsing with lambda calculus*. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*. Association for Computational Linguistics, Prague, Czech Republic, pages 960–967. <http://www.aclweb.org/anthology/P07-1121>.
- Chunyang Xiao, Marc Dymetman, and Claire Gardent. 2016. *Sequence-based structured prediction for semantic parsing*. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 1341–1350. <http://www.aclweb.org/anthology/P16-1127>.

- Wen-tau Yih, Ming-Wei Chang, Xiaodong He, and Jianfeng Gao. 2015. [Semantic parsing via staged query graph generation: Question answering with knowledge base](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Beijing, China, pages 1321–1331. <http://www.aclweb.org/anthology/P15-1128>.
- Wen-tau Yih, Matthew Richardson, Chris Meek, Ming-Wei Chang, and Jina Suh. 2016. [The value of semantic parse labeling for knowledge base question answering](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Association for Computational Linguistics, Berlin, Germany, pages 201–206. <http://anthology.aclweb.org/P16-2033>.
- John M. Zelle and Raymond J. Mooney. 1996. [Learning to parse database queries using inductive logic programming](#). In *AAAI/IAAI*. AAAI Press/MIT Press, Portland, OR, pages 1050–1055. <http://www.cs.utexas.edu/users/ai-lab/?zelle:aaai96>.
- Luke Zettlemoyer and Michael Collins. 2007. [Online learning of relaxed CCG grammars for parsing to logical form](#). In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*. Association for Computational Linguistics, Prague, Czech Republic, pages 678–687. <http://www.aclweb.org/anthology/D/D07/D07-1071>.
- Luke S. Zettlemoyer and Michael Collins. 2005. [Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars](#). In *UAI '05, Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, Edinburgh, Scotland, July 26-29, 2005*. pages 658–666.
- Kai Zhao, Hany Hassan, and Michael Auli. 2015. [Learning translation models from monolingual continuous representations](#). In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Denver, Colorado, pages 1527–1536. <http://www.aclweb.org/anthology/N15-1176>.