Generative Knowledge Graph Construction: A Review

Hongbin Ye[1,2], **Ningyu Zhang**[1,2] *, **Hui Chen** [3], **Huajun Chen**[1,2]
[1] Zhejiang University & AZFT Joint Lab for Knowledge Engine
[2] Hangzhou Innovation Center, Zhejiang University
[3] Alibaba Group
{yehongbin,zhangningyu,huajunsir}@zju.edu.cn,weidu.ch@alibaba-inc.com

Abstract

Generative Knowledge Graph Construction (KGC) refers to those methods that leverage the sequence-to-sequence framework for building knowledge graphs, which is flexible and can be adapted to widespread tasks. In this study, we summarize the recent compelling progress in generative knowledge graph construction. We present the advantages and weaknesses of each paradigm in terms of different generation targets and provide theoretical insight and empirical analysis. Based on the review, we suggest promising research directions for the future. Our contributions are threefold: (1) We present a detailed, complete taxonomy for the generative KGC methods; (2) We provide a theoretical and empirical analysis of the generative KGC methods; (3) We propose several research directions that can be developed in the future.

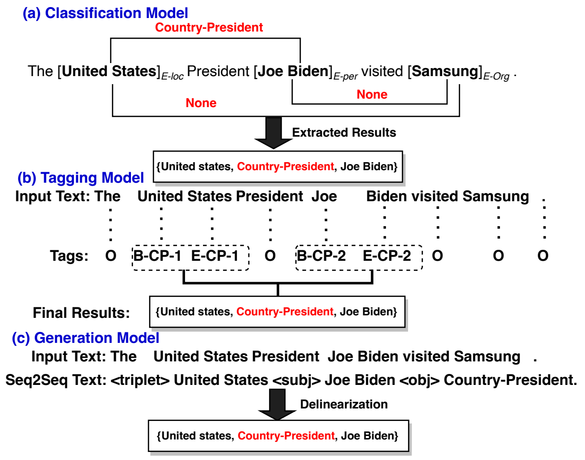

Figure 1: Discrimination and generation methodologies for relation extraction. “Country-President” is the relation, and “CP” is short for “Country-President.”

1 Introduction

Knowledge Graphs (KGs) as a form of structured knowledge have drawn significant attention from academia and the industry (Ji et al., 2022). However, high-quality KGs rely almost exclusively on human-curated structured or semi-structured data. To this end, Knowledge Graph Construction (KGC) is proposed, which is the process of populating (or building from scratch) a KG with new knowledge elements (e.g., entities, relations, events). Conventionally, KGC is solved by employing task-specific discriminators for the various types of information in a pipeline manner (Angeli et al., 2015; Luan et al., 2018; de Sá Mesquita et al., 2019; Zhang et al., 2022a), typically including (1) entity discovery or named entity recognition (Sang and Meulder, 2003), (2) entity linking (Milne and Witten, 2008), (3) relation extraction (Zelenko et al., 2003) and (4) event extraction (Du and Cardie, 2020). However, this presents limitations of error population and poor adaptability for different tasks.

* Corresponding author.

Generative Knowledge Graph Construction. Some generative KGC methods based on the sequence-to-sequence (Seq2Seq) framework are proposed to overcome this barrier. Early work (Zeng et al., 2018) has explored using the generative paradigm to solve different entity and relation extraction tasks. Powered by fast advances of generative pre-training such as T5 (Raffel et al., 2020), and BART (Lewis et al., 2020), Seq2Seq paradigm has shown its great potential in unifying widespread NLP tasks. Hence, more generative KGC works (Yan et al., 2021a; Paolini et al., 2021; Lu et al., 2022) have been proposed, showing appealing performance in benchmark datasets. Figure 1 illustrates an example of generative KGC for relation extraction. The target triple is preceded by the tag <triple>, and the head entity, tail entity, and relations are also specially tagged, allowing the structural knowledge (corresponding to the output) to be obtained by inverse linearization. Despite the success of numerous generative KGC approaches, these works scattered among various tasks have not been systematically reviewed and analyzed.

Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, pages 1–17
December 7-11, 2022 ©2022 Association for Computational Linguistics

Present work In this paper, we summarize recent progress in generative KGC (An timeline of generative KGC can be found in Appendix A) and maintain a public repository for research convenience[1]. We propose to organize relevant work by the generation target of models and also present the axis of the task level (Figure 3):

- **Comprehensive review with new taxonomies**. We conduct the **first** comprehensive review of generative KGC together with new taxonomies. We review the research with different generation targets for KGC with a comprehensive comparison and summary (§3).

- **Theoretical insight and empirical analysis**. We provide in-depth theoretical and empirical analysis for typical generative KGC methods, illustrating the advantages and disadvantageous of different methodologies as well as remaining issues (§4).

- **Wide coverage on emerging advances and outlook on future directions**. We provide comprehensive coverage of emerging areas, including prompt-based learning. This review provides a summary of generative KGC and highlights future research directions (§5).

Related work As this topic is relatively nascent, only a few surveys exist. Closest to our work, Ji et al. (2022) covers methods for knowledge graph construction, representation learning, and applications, which mainly focus on general methods for KGC. Zhu et al. (2022) provides a systematic survey for multi-modal knowledge graph construction and review the challenges, progress, and opportunities. For general NLP, Min et al. (2021) survey recent work that uses these large language models to solve tasks via text generation approaches, which has overlaps in generation methodologies for information extraction. Different from those surveys, in this paper, we conduct a literature review on generative KGC, hoping to systematically understand the methodologies, compare different methods and inspire new ideas.

[1] `https://github.com/zjunlp/Generative_KG_Construction_Papers`

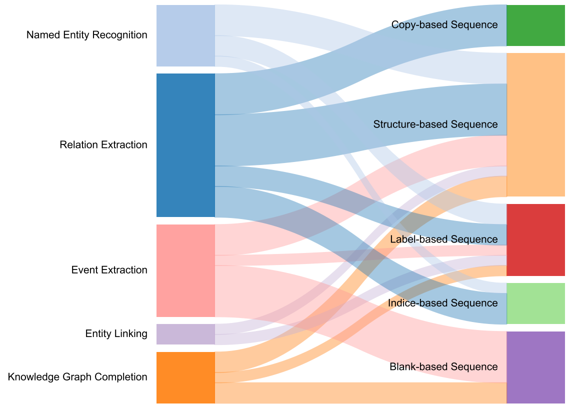

Figure 2: Sankey diagram of knowledge graph construction tasks with different generative paradigms.

2 Preliminary on Knowledge Graph Construction

2.1 Knowledge Graph Construction

Knowledge Graph Construction mainly aims to extract structural information from unstructured texts, such as Named Entity Recognition (NER) (Chiu and Nichols, 2016), Relation Extraction (RE) (Zeng et al., 2015), Event Extraction (EE) (Chen et al., 2015), Entity Linking (EL) (Shen et al., 2015), and Knowledge Graph Completion (Lin et al., 2015).

Generally, KGC can be regarded as structure prediction tasks, where a model is trained to approximate a target function $F(x) \to y$, where $x \in \mathcal{X}$ denotes the input data and $y \in \mathcal{Y}$ denotes the output structure sequence. For instance, given a sentence, "*Steve Jobs and Steve Wozniak co-founded Apple in 1977.*":

Named Entity Recognition aims to identify the types of entities, *e.g.*, '*Steve Job*', '*Steve Wozniak*' $\Rightarrow$ `PERSON`, '*Apple*' $\Rightarrow$ `ORG`;

Relation Extraction aims to identify the relationship of the given entity pair $\langle$*Steve Job*, *Apple*$\rangle$ as `founder`;

Event Extraction aims to identify the event type as `Business Start-Org` where '*co-founded*' triggers the event and (*Steve Jobs*, *Steve Wozniak*) are participants in the event as `AGENT` and `Apple` as `ORG` respectively.

Entity Linking aims to link the mention *Steve Job* to `Steven Jobs (Q19837)` on Wikidata, and *Apple* to `Apple (Q312)` as well.

Knowledge Graph Completion aims to complete incomplete triples $\langle$*Steve Job*, *create*, *?*$\rangle$ for blank entities `Apple`, `NeXT Inc.` and `Pixar`.

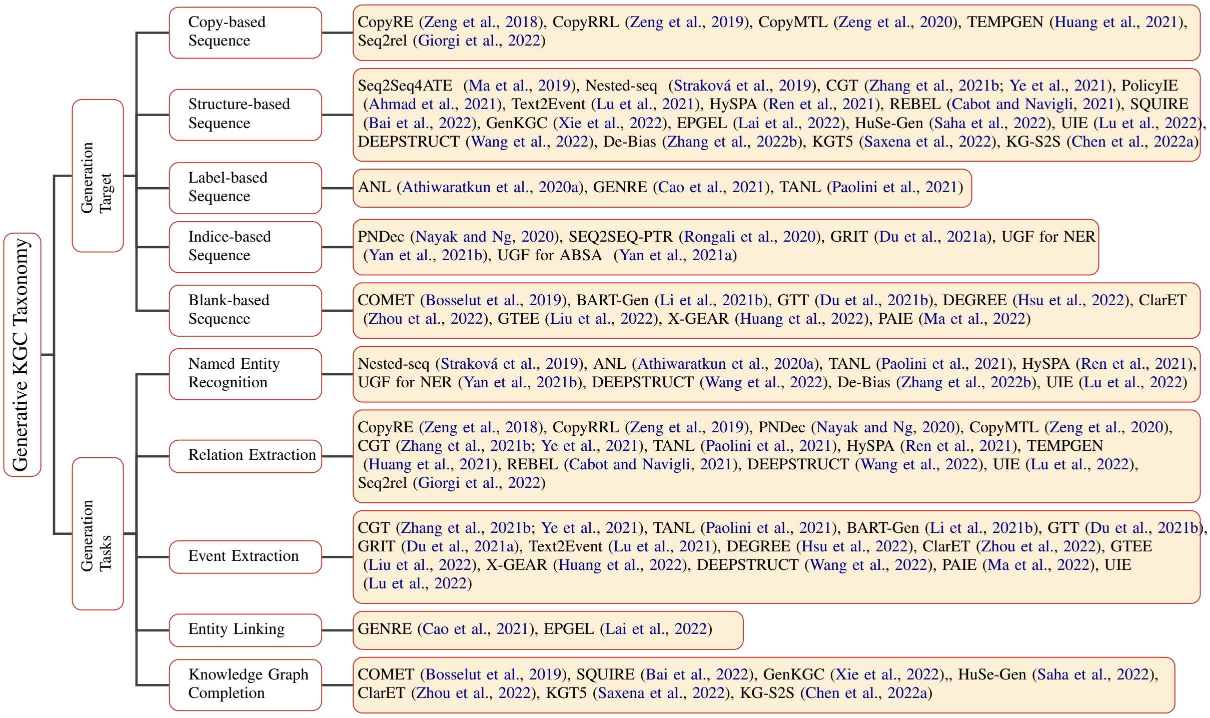

Figure 3: Taxonomy of Generative Knowledge Graph Construction.

2.2 Discrimination and Generation Methodologies

In this section, we introduce the background of discrimination and generation methodologies for KGC. The goal of the discrimination model is to predict the possible label based on the characteristics of the input sentence. As shown in Figure 1, given annotated sentence x and a set of potentially overlapping triples $t_j = \{(s, r, o)\}$ in x, we aim to maximize the data likelihood during the training process:

$$p_{cls}(t|x) = \prod_{(s,r,o) \in t_j} p\left((s, r, o) \mid x_j\right) \tag{1}$$

Another method of discrimination is to output tags using sequential tagging for each position i (Zheng et al., 2017; Dai et al., 2019; Yu et al., 2020; Li et al., 2020b; Liu et al., 2021a). As shown in Figure 1, for an n-word sentence x, n different tag sequences are annotated based on "BIESO" (Begin, Inside, End, Single, Outside) notation schema. The size of a set of pre-defined relations is $|R|$, and the related role orders are represented by "1" and "2". During the training model, we maximize the log-likelihood of the target tag sequence using the hidden vector h_i at each position i:

$$p_{tag}(y \mid x) = \frac{\exp(h_i, y_i)}{\sum_{\mathbf{y}' \in R} \exp\left(\exp(h_i, y_i')\right)} \tag{2}$$

For the generation model, if x is the input sentence and y the result of linearized triplets, the target for the generation model is to autoregressively generate y given x:

$$p_{gen}(y \mid x) = \prod_{i=1}^{\operatorname{len}(y)} p_{gen}\left(y_i \mid y_{<i}, x\right) \tag{3}$$

By fine-tuning seq2seq model (e.g. MASS (Song et al., 2019) , T5 (Raffel et al., 2020) , and BART (Lewis et al., 2020)) on such a task, using the cross-entropy loss, we can maximize the log-likelihood of the generated linearized triplets.

2.3 Advantages of the Generation Methods

While the previous discriminative methods (Wei et al., 2020; Shang et al., 2022) extracts relational triples from unstructured text according to a pre-defined schema to efficiently construct large-scale knowledge graphs, these elaborate models focus on solving a specific task of KGC, such as predicting relation and event information from a segment of input text which often requires multiple models to process. The idea of formulating KGC tasks as sequence-to-sequence problems (Lu et al., 2022) will be of great benefit to develop a universal architecture to solve different tasks, which can be free from the constraints of dedicated architectures, isolated models, and specialized knowledge sources.

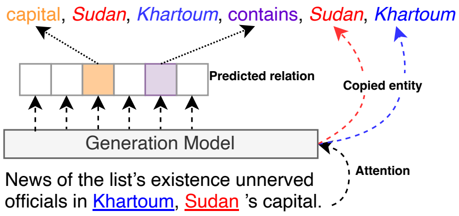

Figure 4: Copy-based Sequence.

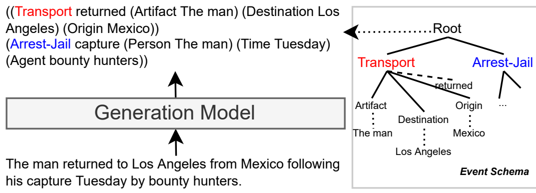

Figure 5: Structure-linearized Sequence.

In addition, generative models can be pre-trained in multiple downstream tasks by structurally consistent linearization of the text, which facilitates the transition from traditional understanding to structured understanding and increases knowledge sharing (Wang et al., 2022). In contexts with nested labels in NER (Straková et al., 2019), the proposed generative method implicitly models the structure between named entities, thus avoiding the complex multi-label mapping. Extracting overlapping triples in RE is also difficult to handle for traditional discriminative models, Zeng et al. (2018) introduce a fresh perspective to revisit the RE task with a general generative framework that addresses the problem by end-to-end model. In short, new directions can be explored for some hard-to-solve problems through paradigm shifts.

Note that the discriminative and generative methods are not simply superior or inferior due to the proliferation of related studies. The aim of this paper is to summarize the characteristics of different generative paradigms in KGC tasks and provide a promising perspective for future research.

3 Taxonomy of Generative Knowledge Graph Construction

In this paper, we mainly consider the following five paradigms that are widely used in KGC tasks based on generation target, i.e. *copy-based Sequence*, *structure-linearized Sequence*, *label-augmented Sequence*, *indice-based Sequence*, and *blank-based Sequence*. As shown in Figure 2, these paradigms have demonstrated strong dominance in many mainstream KGC tasks. In the following sections, we introduce each paradigm as shown in Figure 3.

3.1 Copy-based Sequence

This paradigm refers to developing more robust models to copy the corresponding token (entity) directly from the input sentence during the generation process. Zeng et al. (2018) designs an end-to-end model based on a copy mechanism to solve the triple overlapping problem. As shown in Figure 4, the model copies the head entity from the input sentence and then the tail entity. Similarly, relations are generated from target vocabulary, which is restricted to the set of special relation tokens. This paradigm avoids models generating ambiguous or hallucinative entities. In order to identify a reasonable triple extraction order, Zeng et al. (2019) converts the triplet generation process into a reinforcement learning process, enabling the copy mechanism to follow an efficient generative order. Since the entity copy mechanism relies on unnatural masks to distinguish between head and tail entities, Zeng et al. (2020) maps the head and tail entities to fused feature space for entity replication by an additional nonlinear layer, which strengthens the stability of the mechanism. For document-level extraction, Huang et al. (2021) proposes a TOP-k copy mechanism to alleviate the computational complexity of entity pairs.

3.2 Structure-linearized Sequence

This paradigm refers to utilizing structural knowledge and label semantics, making it prone to handling a unified output format. Lu et al. (2021) proposes an end-to-end event extraction model based on T5, where the output is a linearization of the extracted knowledge structure as shown in Figure 5. In order to avoid introducing noise, it utilizes the event schema to constrain decoding space, ensuring the output text is semantically and structurally legitimate. Lou et al. (2021) reformulates event detection as a Seq2Seq task and proposes a Multi-Layer Bidirectional Network (MLBiNet) to capture the document-level association of events and semantic information simultaneously. Besides, Zhang et al. (2021b); Ye et al. (2021) introduce a contrastive learning framework with a batch dynamic attention masking mechanism to overcome the contradiction in meaning that generative architectures may produce unreliable sequences (Zhu et al., 2020). Similarly, Cabot and Navigli (2021) employs a simple

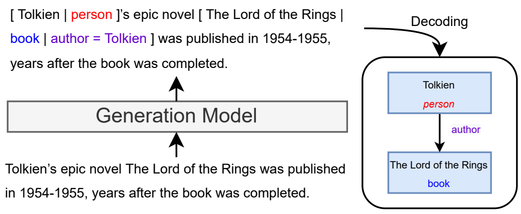

Figure 6: Label-augmented Sequence.

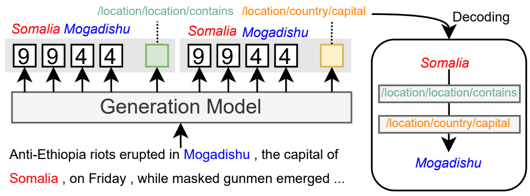

Figure 7: Indice-based Sequence.

triplet decomposition method for the relation extraction task, which is flexible and can be adapted to unified domains or longer documents.

In the nested NER task, Straková et al. (2019) proposes a flattened encoding algorithm, which outputs multiple NE tags following the BILOU scheme. The multi-label of a word is a concatenation of all intersecting tags from highest priority to lowest priority. Similarly, Zhang et al. (2022b) eliminates the incorrect biases in the generation process according to the theory of backdoor adjustment. In EL task, Cao et al. (2021) proposes Generative ENtity REtrieval (GENRE) in an autoregressive fashion conditioned on the context, which captures fine-grained interactions between context and entity name. Moreover, Wang et al. (2022); Lu et al. (2022) extends the domain to structural heterogeneous information extraction by proposing a unified task-agnostic generation framework.

3.3 Label-augmented Sequence

This paradigm refers to utilizing the extra markers to indicate specific entities or relationships. As shown in Figure 6, Athiwaratkun et al. (2020b) investigates the label-augmented paradigm for various structure prediction tasks. The output sequence copies all words in the input sentence, as it helps to reduce ambiguity. In addition, this paradigm uses square brackets or other identifiers to specify the tagging sequence for the entity of interest. The relevant labels are separated by the separator "|" within the enclosed brackets. Meanwhile, the labeled words are described with natural words so that the potential knowledge of the pre-trained model can be leveraged (Paolini et al., 2021). Similarly, Athiwaratkun et al. (2020a) naturally combines tag semantics and shares knowledge across multiple sequence labeling tasks. To retrieve entities by generating their unique names, Cao et al. (2021) extends the autoregressive framework to capture the relations between context and entity name by effectively cross-encoding both. Since the length of the gold decoder targets is often longer than the corresponding input length, this paradigm is unsuitable for document-level tasks because a great portion of the gold labels will be skipped.

3.4 Indice-based Sequence

This paradigm generates the indices of the words in the input text of interest directly and encodes class labels as label indices. As the output is strictly restricted, it will not generate indices that corresponding entities do not exist in the input text, except for relation labels. Nayak and Ng (2020) apply the method to the relation extraction task, enabling the decoder to find all overlapping tuples with full entity names of different lengths. As shown in Figure 7, given the input sequence x, the output sequence y is generated via the indices: $y = [b_1, e_1, t_1, \dots, b_i, e_i, t_i, \dots, b_k, e_k, t_k]$ where b_i and e_i indicates the begin and end indices of a entity tuple, t_i is the index of the entity type, and k is the number of entity tuples. The hidden vector is computed at decoding time by the pointer network (Vinyals et al., 2015) to get the representation of the tuple indices. Besides, Yan et al. (2021b) explores the idea of generating indices for NER, which can be applied to different settings such as flat, nested, and discontinuous NER. In addition, Du et al. (2021a) applies the method to a role-filler entity extraction task by implicitly capturing noun phrase coreference structure.

3.5 Blank-based Sequence

This paradigm refers to utilizing templates to define the appropriate order and relationship for the generated spans. Du et al. (2021b) explores a blank-based form for event extraction tasks which includes special tokens representing event information such as event types. Li et al. (2021b) frames document-level event argument extraction as con-

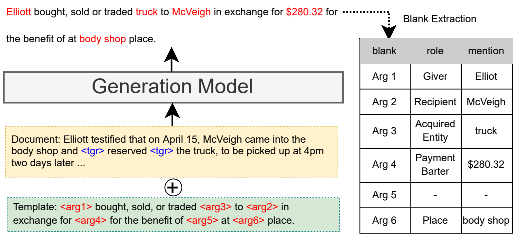

Figure 8: Blank-based Sequence.

ditional generation given a template and introduces the new document-level informative to aid the generation process. As shown in Figure 8, the template refers to a text describing an event type, which adds blank argument role placeholders. The output sequences are sentences where the blank placeholders are replaced by specific event arguments. Besides, Hsu et al. (2022) focuses on low-resource event extraction and proposes a data-efficient model called DEGREE, which utilizes label semantic information. Huang et al. (2022) designs a language-agnostic template to represent the event argument structures, which facilitate the cross-lingual transfer. Instead of conventional heuristic threshold tuning, Ma et al. (2022) proposes an effective yet efficient model PAIE for extracting multiple arguments with the same role.

3.6 Comparison and Discussion

Recently, the literature on generative KGC has been growing rapidly. A unifying theme across many of these methods is that of end-to-end architecture or the idea that the knowledge extraction can be redefined as *text sequence to structure generation* task. Generative models can decode and control extraction targets on demand for different specific tasks, scenarios, and settings (i.e., different schema). However, due to the different forms of specific KGC tasks, there is still some disagreement in the utilization of the generation paradigms.

As shown in Table 1, we make a comprehensive comparison among the paradigms mentioned above via rating based on different evaluation scopes: 1) **Semantic utilization** refers to the degree to which the model leverages the semantics of the labels. In principle, we believe that the closer the output form is to natural language, the smaller the gap between the generative model and the training task. We observe that the blank-based paradigm has a clear advantage in this scope, which uses manually constructed templates to make the output close to natural language fluency. 2) **Search space** refers to the vocabulary space searched by the decoder. Due to the application of the constraint decoding mechanism, some structure-based methods can be reduced to the same decoding space as the copy-based methods. In addition, the indice-based paradigm uses a pointer mechanism that constrains the output space to the length of the input sequence. 3) **Application scope** refers to the range of KGC tasks that can be applied. We believe that architectures with the ability to organize information more flexibly have excellent cross-task migration capabilities such as structure-based, label-based and blank-based paradigms. 4) **Template cost** refers to the cost of constructing the input and golden output text. We observe that most paradigms do not require complex template design and rely only on linear concatenation to meet the task requirement. However, the blank-based paradigm requires more labor consumption to make the template conform to the semantic fluency requirement.

Totally in line with recent trends in NLP, a growing number of unified generation strategies require more universal architectures (Deng et al., 2021; Li et al., 2021a), as they allow a remarkable degree of output flexibility. We think that future research should focus on unifying cross-task models and further improving decoding efficiency.

4 Analysis

4.1 Theoretical Insight

This section provides theoretical insight into optimization and inference for generative KGC. For optimization, NLG are normally modeled by parameterized probabilistic models p_{gen} over text strings $\mathbf{y} = \langle y_1, y_2, \ldots \rangle$ decomposed by words y_t:

$$p_{gen}(y \mid x) = \prod_{i=1}^{\text{len}(y)} p_{gen}\left(y_i \mid y_{<i}, x\right) \tag{4}$$

where y consists of all possible strings that can be constructed from words in the model's vocabulary $\mathcal{V}$. Note that the output y can take on a variety of forms depending on the task, e.g., entities, relational triples, or an event structure. Usually, the model will limit the target set by pre-defined schema as $\mathcal{Y}_{\mathcal{T}} \subset \mathcal{Y}$. The optimization procedure will be taken to estimate the parameters with log-likelihood maximization as follows:

$$L(\boldsymbol{\theta}; \mathcal{T}) = -\sum_{\mathbf{y} \in \mathcal{T}} \log q(\mathbf{y}) \tag{5}$$

Taxonomy	Generative Strategy	Representative Model	Evaluation Scope			
			SU↑	SS↓	AS↓	TS↓
Copy-based (§ 3.1)	Directly copy entity	CopyRE (Zeng et al., 2018)	L	L	M	L
	Restricted target vocabulary	Seq2rel (Giorgi et al., 2022)	L	L	H	L
Structure-based (§ 3.2)	Per-token tag encoding	Nested-seq (Straková et al., 2019)	L	L	H	L
	Faithful contrastive learning	CGT (Zhang et al., 2021b)	M	M	H	L
	Prefix tree constraint decoding	TEXT2EVENT (Lu et al., 2021)	M	M	H	L
	Triplet linearization	REBEL (Cabot and Navigli, 2021)	M	H	M	L
	Entity-aware hierarchical decoding	GenKGC (Xie et al., 2022)	M	L	M	L
	Unified structure generation	UIE (Lu et al., 2022)	M	H	H	L
	Reformulating triple prediction	DEEPSTRUCT (Wang et al., 2022)	M	H	H	L
	Query Verbalization	KGT5 (Saxena et al., 2022)	M	H	M	L
Label-based (§ 3.3)	Augmented natural language	TANL (Paolini et al., 2021)	M	H	H	L
Indice-based (§ 3.4)	Pointer mechanism	PNDec (Nayak and Ng, 2020)	L	L	M	L
	Pointer selection	GRIT (Du et al., 2021a)	M	L	M	L
Blank-based (§ 3.5)	Template filling as generation	GTT (Du et al., 2021b)	H	H	H	H
	Prompt semantic guidance	DEGREE (Hsu et al., 2022)	H	H	H	H
	Language-agnostic template	X-GEAR (Huang et al., 2022)	H	M	H	H

Table 1: Comparison of generation methods from different evaluation scopes. "SU" indicates semantic utilization, "SS" indicates search space, "AS" indicates application scope, and "TS" indicates template cost. We divide the degree into three grades: L (low), M (middle), and H (high), and the ↑ indicates that the higher grade performance is better while the ↓ is the opposite.

where $\boldsymbol{\theta}$ are the model parameters. Notably, with small output space (e.g., methods with the indice-based sequence in §3.4), the model can converge faster. However, the model with a small output space may fail to utilize rich semantic information from labels or text (like models in §3.5). In short, the design of output space is vital for generative KGC, and it is necessary to balance parametric optimization as well as semantic utilization.

For inference, we argue that sequence decoding in the generation is an essential procedure for generative KGC. Given the probabilistic nature of q, the decoding process will select words that maximize the probability of the resulting string. Vanilla decoding solutions such as beam search or greedy have been investigated in generative KGC. On the one hand, knowledge-guided (or schema-guided) decoding has become the mainstay for many generative KGC tasks. For example, Lu et al. (2021) proposes Text2Event in which words are decoded through a prefix tree based on pre-defined schema. On the other hand, non-autoregressive parallel decoding has also been leveraged for generative KGC. Sui et al. (2021) formulates end-to-end knowledge base population as a direct set generation problem, avoiding considering the order of multiple facts. Note that the decoding mechanism plays a vital role in inference speed and quality. We argue that it is necessary to develop sophisticated, efficient decoding strategies (e.g., with guidance from KG) for generative KGC.

4.2 Empirical Analysis

To investigate the effect of different generation methods, we conduct an analysis of the experimental results of existing generative KGC work. Due to space limitations of the article, we only select two representative tasks of entity/relation extraction and event extraction with NYT and ACE datasets[2]. Table 2 shows the performance of discrimination models and generative models on the NYT datasets. We can observe that: 1) Structure-based and label-based methods both achieve similar extraction performance compared with all discrimination models on NYT datasets. We believe this is because they can better utilize label semantics and structural knowledge than other generation methods. 2) Although the discrimination methods obtain good performance, the performance of the generation methods has been improved more vastly in recent years, so we have reason to believe that they will have greater application scope in the near future. In addition, we also show the performance of the non-autoregressive method on two datasets, and we discuss the promising value of this method in § 5. We observe that parallel generation of the unordered triple set can obtain comparable

[2]Results are taken from existing papers.

Type	Models	NYT		
		P	R	F
Discrimination	CasRel (Wei et al., 2020)	89.7	89.5	89.6
	TPLinker (Wang et al., 2020)	91.4	92.6	92.0
	OneRel (Shang et al., 2022)	92.8	92.9	92.8
Copy-based	CopyRE (Zeng et al., 2018)	61.0	56.6	58.7
	CopyRRL (Zeng et al., 2019)	77.9	67.2	72.1
	CopyMTL (Zeng et al., 2020)	75.7	68.7	72.0
Structure-based	CGT (Ye et al., 2021)	94.7	84.2	89.1
	REBEL (Cabot and Navigli, 2021)	91.5	92.0	91.8
	UIE (Lu et al., 2022)	-	-	93.5
	DEEPSTRUCT (Wang et al., 2022)	-	-	93.9
Label-based	TANL (Paolini et al., 2021)	-	-	90.8
Indice-based	PNDec (Nayak and Ng, 2020)	89.3	78.8	83.8
Others*	SPN (Sui et al., 2020)	93.3	91.7	92.5
	Seq2UMTree (Zhang et al., 2020b)	79.1	75.1	77.1

Table 2: Main results of NYT dataset. The top section refers to the discrimination models, and the bottom section indicates generation models. "*" refers to the non-autoregressive models.

Type	Models	Trigger		Argument	
		Id	Cl	Id	Cl
Discrimination	JMEE (Liu et al., 2018)	75.9	73.7	68.4	60.3
	DYGIE++ (Wadden et al., 2019)	-	69.7	53.0	48.8
	OneIE (Lin et al., 2020)	78.6	75.2	60.7	58.6
	QAEE (Du and Cardie, 2020)	75.8	72.4	55.3	53.3
	MQAEE (Li et al., 2020a)	74.5	71.7	55.2	53.4
	RCEE (Liu et al., 2020)	-	74.9	-	63.6
Structure-based	TEXT2EVENT (Lu et al., 2021)	-	71.9	-	53.8
	UIE (Lu et al., 2022)	-	73.4	-	54.8
	DEEPSTRUCT (Wang et al., 2022)	73.5	69.8	59.4	56.2
Label-based	TANL (Paolini et al., 2021)	72.9	68.4	50.1	47.6
Blank-based	BART-Gen (Du et al., 2021b)	74.4	71.1	55.2	53.7
	DEGREE (Hsu et al., 2022)	-	73.3	-	55.8
	GTEE (Liu et al., 2022)	-	72.6	-	55.8
	PAIE (Ma et al., 2022)	-	-	75.7*	72.7*

Table 3: F1 results (%) of ACE-2005. The top section refers to the discrimination models, and the bottom section indicates the generation models. Id is Identification, and Cl is Classification. "*" refers to experiments only in argument extraction tasks with the golden trigger.

performance with advanced discriminative models, noting that non-autoregressive methods have better decoding efficiency and training efficiency.

From Table 3, we observe that generation methods can obtain comparable performance compared with discrimination models on event extraction tasks. Since the framework of event extraction has a hierarchical structure (i.e., it is usually decomposed into two subtasks: trigger extraction and argument extraction), structure-based methods have a supervised learning framework for the sequence-to-structure generation, while schema constraints guarantee structural and semantic legitimacy. In addition, owing to the complete template design of the Blank-based approach, PLMs can understand complex task knowledge, structural knowledge of the extraction framework, and label semantics in a natural language manner.

5 Future Directions

Though lots of technical solutions have been proposed for generative KGC as surveyed, there remain some potential directions:

Generation Architecture. Most of the recent generative KGC frameworks face serious homogenization with Transformer. For enhancing interpretability, we argue that neuro-symbolic models (i.e., a reasoning system that integrates neural and symbolic) (Zhang et al., 2021a; Galassi, 2021; Negro and Pons, 2022) can be designed for generative

KGC. In addition, some cutting-edge technologies such as spiking neural network (Tavanaei et al., 2019), dynamic neural networks (Xu and McAuley, 2022), ordinary differential equations (Li et al., 2022a) and diffusion models (Dhariwal and Nichol, 2021) can also provide promising architectures.

Generation Quality. Considering the target reliability of generation methods, more sophisticated strategies can be leveraged to control the quality of generative KGC, including: 1) Control code construction (Keskar et al., 2019; Dou et al., 2021); 2) Decoding strategy such as introducing external feedback (Holtzman et al., 2018) and generative discriminator (Krause et al., 2021); 3) Loss function design (Chan et al., 2021); 4) Prompt design (Brown et al., 2020; Qian et al., 2022); 5) Retrieval augmentation (Li et al., 2022b); 6) Write-then-Edit strategy (Dathathri et al., 2020); 7) Diffusion process (Li et al., 2022c; Gong et al., 2022).

Training Efficiency. In practical applications, it is essential to reduce data annotation and training costs. One idea is to freeze most of the generation model parameters (Liu et al., 2021b; Li and Liang, 2021; Chen et al., 2022b) or leverage prompt learning (Chen et al., 2022e). Another idea is that knowledge decoupling intervention training models can reduce parameter redundancy (Wang et al., 2021; Borgeaud et al., 2021; Khandelwal et al., 2020; Chen et al., 2022d,c).

Universal Deployment. Inspired by the T5 (Raffel et al., 2020), which transforms all NLP tasks into Text-to-Text tasks, generation models can be generalized to the multi-task and multi-modal domain. Therefore, instead of improvements being prone to be exclusive to a single task, domain, or dataset, we argue that it is beneficial to study the framework to advocate for a unified view of KGC, such as the wonderful work UIE (Lu et al., 2022). Furthermore, it is efficient for real-world deployment when we can provide a single model to support widespread KGC tasks (Zhang et al., 2020a).

Inference Speed. To be noted, although previous work has treated KGC as end-to-end generative tasks, they are still limited by auto-regressive decoders. However, the autoregressive decoder generates each token based on previously generated tokens during inference, and this process is not parallelizable. Therefore, it is beneficial to develop a fast inference model for generative KGC. Previously, Sui et al. (2020) utilizes the transformer-based non-autoregressive decoder (Gu et al., 2018) as a triple set generator that can predict all triples at once. Sui et al. (2021) also formulates end-to-end knowledge base population as a direct set generation problem. Zhang et al. (2020b) proposes a two-dimensional unordered multitree allowing prediction deviations not to aggregate and affect other triples. To sum up, the non-autoregressive approach applied to KGC proves to be effective in solving the exposure bias and overfitting problems. Likewise, the semi-autoregressive decoding (Wang et al., 2018) preserves the autoregressive approach within the block to ensure consistency while improving the tuple output efficiency. Additionally, pathways (Barham et al., 2022) can dynamically assign competencies to different parts of the neural network, which is faster and more efficient as it does not activate the entire network for each task.

6 Conclusion and Vision

In this paper, we provide an overview of generative KGC with new taxonomy, theoretical insight and empirical analysis, and several research directions. Note that the generative paradigm for KGC has the potential advantages of unifying different tasks and better utilizing semantic information. In the future, we envision a more potent synergy between the methodologies from the NLG and knowledge graph communities. We hope sophisticated and efficient text generation models to be increasingly contributed to improving the KGC performance. On the converse, we expect symbolic structure in KG can have potential guidance for text generation.

7 Limitations

In this study, we provide a review of generative KGC. Due to the page limit, we cannot afford the technical details for models. Moreover, we only review the works within five years, mainly from the ACL, EMNLP, NAACL, COLING, AAAI, IJCAI, etc. We will continue adding more related works with more detailed analysis.

Acknowledgment

We want to express gratitude to the anonymous reviewers. This work was supported by the National Natural Science Foundation of China (No.62206246, 91846204 and U19B2027), Zhejiang Provincial Natural Science Foundation of China (No. LGG22F030011), Ningbo Natural Science Foundation (2021J190), and Yongjiang Talent Introduction Programme (2021A-156-G).

References

Wasi Uddin Ahmad, Jianfeng Chi, Tu Le, Thomas Norton, Yuan Tian, and Kai-Wei Chang. 2021. Intent classification and slot filling for privacy policies. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 4402–4417. Association for Computational Linguistics.

Gabor Angeli, Victor Zhong, Danqi Chen, Arun Tejasvi Chaganty, Jason Bolton, Melvin Jose Johnson Premkumar, Panupong Pasupat, Sonal Gupta, and Christopher D. Manning. 2015. Bootstrapped self training for knowledge base population. In *Proceedings of the 2015 Text Analysis Conference, TAC 2015, Gaithersburg, Maryland, USA, November 16-17, 2015, 2015*. NIST.

Ben Athiwaratkun, Cícero Nogueira dos Santos, Jason Krone, and Bing Xiang. 2020a. Augmented natural language for generative sequence labeling. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 375–385. Association for Computational Linguistics.

Ben Athiwaratkun, Cicero Nogueira dos Santos, Jason Krone, and Bing Xiang. 2020b. Augmented natural language for generative sequence labeling. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Yushi Bai, Xin Lv, Juanzi Li, Lei Hou, Yincen Qu, Zelin Dai, and Feiyu Xiong. 2022. SQUIRE: A sequence-to-sequence framework for multi-hop knowledge graph reasoning. *CoRR*, abs/2201.06206.

Paul Barham, Aakanksha Chowdhery, Jeff Dean, Sanjay Ghemawat, Steven Hand, Dan Hurt, Michael Isard, Hyeontaek Lim, Ruoming Pang, Sudip Roy, Brennan Saeta, Parker Schuh, Ryan Sepassi, Laurent El Shafey, Chandramohan A. Thekkath, and Yonghui Wu. 2022. Pathways: Asynchronous distributed dataflow for ML. In *Proceedings of Machine Learning and Systems 2022, MLSys 2022, Santa Clara, CA, USA, August 29 - September 1, 2022*. mlsys.org.

Sebastian Borgeaud, Arthur Mensch, Jordan Hoffmann, Trevor Cai, Eliza Rutherford, Katie Millican, George van den Driessche, Jean-Baptiste Lespiau, Bogdan Damoc, Aidan Clark, Diego de Las Casas, Aurelia Guy, Jacob Menick, Roman Ring, Tom Hennigan, Saffron Huang, Loren Maggiore, Chris Jones, Albin Cassirer, Andy Brock, Michela Paganini, Geoffrey Irving, Oriol Vinyals, Simon Osindero, Karen Simonyan, Jack W. Rae, Erich Elsen, and Laurent Sifre. 2021. Improving language models by retrieving from trillions of tokens. *CoRR*, abs/2112.04426.

Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. 2019. COMET: commonsense transformers for automatic knowledge graph construction. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 4762–4779. Association for Computational Linguistics.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Pere-Lluís Huguet Cabot and Roberto Navigli. 2021. REBEL: relation extraction by end-to-end language generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 16-20 November, 2021*, pages 2370–2381. Association for Computational Linguistics.

Nicola De Cao, Gautier Izacard, Sebastian Riedel, and Fabio Petroni. 2021. Autoregressive entity retrieval. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Alvin Chan, Yew-Soon Ong, Bill Pung, Aston Zhang, and Jie Fu. 2021. Cocon: A self-supervised approach for controlled text generation. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Chen Chen, Yufei Wang, Bing Li, and Kwok-Yan Lam. 2022a. Knowledge is flat: A seq2seq generative framework for various knowledge graph completion. In *Proceedings of the 29th International Conference on Computational Linguistics, COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*, pages 4005–4017. International Committee on Computational Linguistics.

Xiang Chen, Lei Li, Shumin Deng, Chuanqi Tan, Changliang Xu, Fei Huang, Luo Si, Huajun Chen, and Ningyu Zhang. 2022b. LightNER: A lightweight tuning paradigm for low-resource NER via pluggable prompting. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 2374–2387, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Xiang Chen, Lei Li, Ningyu Zhang, Xiaozhuan Liang, Shumin Deng, Chuanqi Tan, Fei Huang, Luo Si, and Huajun Chen. 2022c. Decoupling knowledge from memorization: Retrieval-augmented prompt learning. In *Proceedings of NeurIPS 2022*.

Xiang Chen, Lei Li, Ningyu Zhang, Chuanqi Tan, Fei Huang, Luo Si, and Huajun Chen. 2022d. Relation extraction as open-book examination: Retrieval-enhanced prompt tuning. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*, pages 2443–2448. ACM.

Xiang Chen, Ningyu Zhang, Xin Xie, Shumin Deng, Yunzhi Yao, Chuanqi Tan, Fei Huang, Luo Si, and Huajun Chen. 2022e. Knowprompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction. In *WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022*, pages 2778–2788. ACM.

Yubo Chen, Liheng Xu, Kang Liu, Daojian Zeng, and Jun Zhao. 2015. Event extraction via dynamic multi-pooling convolutional neural networks. In *ACL (1)*, pages 167–176.

Jason P. C. Chiu and Eric Nichols. 2016. Named entity recognition with bidirectional lstm-cnns. *Trans. Assoc. Comput. Linguistics*, 4:357–370.

Dai Dai, Xinyan Xiao, Yajuan Lyu, Shan Dou, Qiaoqiao She, and Haifeng Wang. 2019. Joint extraction of entities and overlapping relations using position-attentive sequence labeling. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pages 6300–6308. AAAI Press.

Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. 2020. Plug and play language models: A simple approach to controlled text generation. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Filipe de Sá Mesquita, Matteo Cannaviccio, Jordan Schmidek, Paramita Mirza, and Denilson Barbosa. 2019. Knowledgenet: A benchmark dataset for knowledge base population. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pages 749–758. Association for Computational Linguistics.

Mingkai Deng, Bowen Tan, Zhengzhong Liu, Eric P. Xing, and Zhiting Hu. 2021. Compression, transduction, and creation: A unified framework for evaluating natural language generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 7580–7605. Association for Computational Linguistics.

Prafulla Dhariwal and Alexander Quinn Nichol. 2021. Diffusion models beat gans on image synthesis. In *Advances in Neural Information Processing Systems 34: Annual Conference on Neural Information Processing Systems 2021, NeurIPS 2021, December 6-14, 2021, virtual*, pages 8780–8794.

Zi-Yi Dou, Pengfei Liu, Hiroaki Hayashi, Zhengbao Jiang, and Graham Neubig. 2021. Gsum: A general framework for guided neural abstractive summarization. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 4830–4842. Association for Computational Linguistics.

Xinya Du and Claire Cardie. 2020. Event extraction by answering (almost) natural questions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 671–683. Association for Computational Linguistics.

Xinya Du, Alexander M. Rush, and Claire Cardie. 2021a. GRIT: generative role-filler transformers for document-level event entity extraction. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, EACL 2021, Online, April 19 - 23, 2021*, pages 634–644. Association for Computational Linguistics.

Xinya Du, Alexander M. Rush, and Claire Cardie. 2021b. Template filling with generative transformers. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 909–914. Association for Computational Linguistics.

Andrea Galassi. 2021. *Deep Networks and Knowledge: from Rule Learning to Neural-Symbolic Argument Mining*. Ph.D. thesis, University of Bologna, Italy.

John M. Giorgi, Gary D. Bader, and Bo Wang. 2022. A sequence-to-sequence approach for document-level relation extraction. In *Proceedings of the 21st Workshop on Biomedical Language Processing, BioNLP@ACL 2022, Dublin, Ireland, May 26, 2022*, pages 10–25. Association for Computational Linguistics.

Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, and LingPeng Kong. 2022. Diffuseq: Sequence to sequence text generation with diffusion models.

Jiatao Gu, James Bradbury, Caiming Xiong, Victor O. K. Li, and Richard Socher. 2018. Non-autoregressive neural machine translation. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net.

Ari Holtzman, Jan Buys, Maxwell Forbes, Antoine Bosselut, David Golub, and Yejin Choi. 2018. Learning to write with cooperative discriminators. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 1638–1649. Association for Computational Linguistics.

I-Hung Hsu, Kuan-Hao Huang, Elizabeth Boschee, Scott Miller, Prem Natarajan, Kai-Wei Chang, and Nanyun Peng. 2022. Event extraction as natural language generation. *NAACL*, abs/2108.12724.

Kuan-Hao Huang, I-Hung Hsu, Prem Natarajan, Kai-Wei Chang, and Nanyun Peng. 2022. Multilingual generative language models for zero-shot cross-lingual event argument extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 4633–4646. Association for Computational Linguistics.

Kung-Hsiang Huang, Sam Tang, and Nanyun Peng. 2021. Document-level entity-based extraction as template generation. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 5257–5269. Association for Computational Linguistics.

Shaoxiong Ji, Shirui Pan, Erik Cambria, Pekka Marttinen, and Philip S. Yu. 2022. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Trans. Neural Networks Learn. Syst.*, 33(2):494–514.

Nitish Shirish Keskar, Bryan McCann, Lav R. Varshney, Caiming Xiong, and Richard Socher. 2019. CTRL: A conditional transformer language model for controllable generation. *CoRR*, abs/1909.05858.

Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, and Mike Lewis. 2020. Generalization through memorization: Nearest neighbor language models. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net.

Ben Krause, Akhilesh Deepak Gotmare, Bryan McCann, Nitish Shirish Keskar, Shafiq R. Joty, Richard Socher, and Nazneen Fatema Rajani. 2021. Gedi: Generative discriminator guided sequence generation. In *Findings of the Association for Computational Linguistics: EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 16-20 November, 2021*, pages 4929–4952. Association for Computational Linguistics.

Tuan Lai, Heng Ji, and ChengXiang Zhai. 2022. Improving candidate retrieval with entity profile generation for wikidata entity linking. In *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 3696–3711. Association for Computational Linguistics.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 7871–7880. Association for Computational Linguistics.

Bei Li, Quan Du, Tao Zhou, Yi Jing, Shuhan Zhou, Xin Zeng, Tong Xiao, JingBo Zhu, Xuebo Liu, and Min Zhang. 2022a. ODE transformer: An ordinary differential equation-inspired model for sequence generation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 8335–8351. Association for Computational Linguistics.

Fayuan Li, Weihua Peng, Yuguang Chen, Quan Wang, Lu Pan, Yajuan Lyu, and Yong Zhu. 2020a. Event extraction as multi-turn question answering. In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020*, volume EMNLP 2020 of *Findings of ACL*, pages 829–838. Association for Computational Linguistics.

Huayang Li, Yixuan Su, Deng Cai, Yan Wang, and Lemao Liu. 2022b. A survey on retrieval-augmented text generation. *CoRR*, abs/2202.01110.

Juan Li, Ruoxu Wang, Ningyu Zhang, Wen Zhang, Fan Yang, and Huajun Chen. 2020b. Logic-guided semantic representation learning for zero-shot relation classification. In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, pages 2967–2978. International Committee on Computational Linguistics.

Junyi Li, Tianyi Tang, Gaole He, Jinhao Jiang, Xiaoxuan Hu, Puzhao Xie, Zhipeng Chen, Zhuohao Yu, Wayne Xin Zhao, and Ji-Rong Wen. 2021a. Textbox: A unified, modularized, and extensible framework for text generation. In *Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL 2021 - System Demonstrations, Online, August 1-6, 2021*, pages 30–39. Association for Computational Linguistics.

Sha Li, Heng Ji, and Jiawei Han. 2021b. Document-level event argument extraction by conditional generation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 894–908. Association for Computational Linguistics.

Xiang Lisa Li and Percy Liang. 2021. Prefix-tuning: Optimizing continuous prompts for generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 4582–4597. Association for Computational Linguistics.

Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, Percy Liang, and Tatsunori B. Hashimoto. 2022c. Diffusion-lm improves controllable text generation. *CoRR*, abs/2205.14217.

Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. 2015. Learning entity and relation embeddings for knowledge graph completion. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA*, pages 2181–2187. AAAI Press.

Ying Lin, Heng Ji, Fei Huang, and Lingfei Wu. 2020. A joint neural model for information extraction with global features. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 7999–8009. Association for Computational Linguistics.

Jian Liu, Yubo Chen, Kang Liu, Wei Bi, and Xiaojiang Liu. 2020. Event extraction as machine reading comprehension. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 1641–1651. Association for Computational Linguistics.

Kun Liu, Yao Fu, Chuanqi Tan, Mosha Chen, Ningyu Zhang, Songfang Huang, and Sheng Gao. 2021a. Noisy-labeled NER with confidence estimation. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 3437–3445. Association for Computational Linguistics.

Xiao Liu, Heyan Huang, Ge Shi, and Bo Wang. 2022. Dynamic prefix-tuning for generative template-based event extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 5216–5228. Association for Computational Linguistics.

Xiao Liu, Zhunchen Luo, and Heyan Huang. 2018. Jointly multiple events extraction via attention-based graph information aggregation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 1247–1256. Association for Computational Linguistics.

Xiao Liu, Yanan Zheng, Zhengxiao Du, Ming Ding, Yujie Qian, Zhilin Yang, and Jie Tang. 2021b. GPT understands, too. *CoRR*, abs/2103.10385.

Dongfang Lou, Zhilin Liao, Shumin Deng, Ningyu Zhang, and Huajun Chen. 2021. Mlbinet: A cross-sentence collective event detection network. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 4829–4839. Association for Computational Linguistics.

Yaojie Lu, Hongyu Lin, Jin Xu, Xianpei Han, Jialong Tang, Annan Li, Le Sun, Meng Liao, and Shaoyi Chen. 2021. Text2event: Controllable sequence-to-structure generation for end-to-end event extraction. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 2795–2806. Association for Computational Linguistics.

Yaojie Lu, Qing Liu, Dai Dai, Xinyan Xiao, Hongyu Lin, Xianpei Han, Le Sun, and Hua Wu. 2022. Unified structure generation for universal information extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 5755–5772. Association for Computational Linguistics.

Yi Luan, Luheng He, Mari Ostendorf, and Hannaneh Hajishirzi. 2018. Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 3219–3232. Association for Computational Linguistics.

Dehong Ma, Sujian Li, Fangzhao Wu, Xing Xie, and Houfeng Wang. 2019. Exploring sequence-to-sequence learning in aspect term extraction. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 3538–3547. Association for Computational Linguistics.

Yubo Ma, Zehao Wang, Yixin Cao, Mukai Li, Meiqi Chen, Kun Wang, and Jing Shao. 2022. Prompt for extraction? PAIE: prompting argument interaction for event argument extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 6759–6774. Association for Computational Linguistics.

David N. Milne and Ian H. Witten. 2008. Learning to link with wikipedia. In *Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, California, USA, October 26-30, 2008*, pages 509–518. ACM.

Bonan Min, Hayley Ross, Elior Sulem, Amir Pouran Ben Veyseh, Thien Huu Nguyen, Oscar Sainz, Eneko Agirre, Ilana Heintz, and Dan Roth. 2021. Recent advances in natural language processing via large pre-trained language models: A survey. *CoRR*, abs/2111.01243.

Tapas Nayak and Hwee Tou Ng. 2020. Effective modeling of encoder-decoder architecture for joint entity and relation extraction. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 8528–8535. AAAI Press.

Pablo Negro and Claudia Pons. 2022. Artificial intelligence techniques based on the integration of symbolic logic and deep neural networks: A systematic review of the literature. *Inteligencia Artif.*, 25(69):13–41.

Giovanni Paolini, Ben Athiwaratkun, Jason Krone, Jie Ma, Alessandro Achille, Rishita Anubhai, Cícero Nogueira dos Santos, Bing Xiang, and Stefano Soatto. 2021. Structured prediction as translation between augmented natural languages. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net.

Jing Qian, Li Dong, Yelong Shen, Furu Wei, and Weizhu Chen. 2022. Controllable natural language generation with contrastive prefixes. In *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2912–2924. Association for Computational Linguistics.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, 21:140:1–140:67.

Liliang Ren, Chenkai Sun, Heng Ji, and Julia Hockenmaier. 2021. Hyspa: Hybrid span generation for scalable text-to-graph extraction. In *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, Online Event, August 1-6, 2021*, volume ACL/IJCNLP 2021 of *Findings of ACL*, pages 4066–4078. Association for Computational Linguistics.

Subendhu Rongali, Luca Soldaini, Emilio Monti, and Wael Hamza. 2020. Don't parse, generate! A sequence to sequence architecture for task-oriented semantic parsing. In *WWW '20: The Web Conference 2020, Taipei, Taiwan, April 20-24, 2020*, pages 2962–2968. ACM / IW3C2.

Swarnadeep Saha, Prateek Yadav, and Mohit Bansal. 2022. Explanation graph generation via pre-trained language models: An empirical study with contrastive learning. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 1190–1208. Association for Computational Linguistics.

Erik F. Tjong Kim Sang and Fien De Meulder. 2003. Introduction to the conll-2003 shared task: Language-independent named entity recognition. In *Proceedings of the Seventh Conference on Natural Language Learning, CoNLL 2003, Held in cooperation with HLT-NAACL 2003, Edmonton, Canada, May 31 - June 1, 2003*, pages 142–147. ACL.

Apoorv Saxena, Adrian Kochsiek, and Rainer Gemulla. 2022. Sequence-to-sequence knowledge graph completion and question answering. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2814–2828. Association for Computational Linguistics.

Yuming Shang, Heyan Huang, and Xian-Ling Mao. 2022. Onerel: Joint entity and relation extraction with one module in one step. *CoRR*, abs/2203.05412.

Wei Shen, Jianyong Wang, and Jiawei Han. 2015. Entity linking with a knowledge base: Issues, techniques, and solutions. *IEEE Trans. Knowl. Data Eng.*, 27(2):443–460.

Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, and Tie-Yan Liu. 2019. MASS: masked sequence to sequence pre-training for language generation. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research*, pages 5926–5936. PMLR.

Jana Straková, Milan Straka, and Jan Hajic. 2019. Neural architectures for nested NER through linearization. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 5326–5331. Association for Computational Linguistics.

Dianbo Sui, Yubo Chen, Kang Liu, Jun Zhao, Xiangrong Zeng, and Shengping Liu. 2020. Joint entity and relation extraction with set prediction networks. *CoRR*, abs/2011.01675.

Dianbo Sui, Chenhao Wang, Yubo Chen, Kang Liu, Jun Zhao, and Wei Bi. 2021. Set generation networks for end-to-end knowledge base population. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 9650–9660. Association for Computational Linguistics.

Amirhossein Tavanaei, Masoud Ghodrati, Saeed Reza Kheradpisheh, Timothée Masquelier, and Anthony Maida. 2019. Deep learning in spiking neural networks. *Neural Networks*, 111:47–63.

Oriol Vinyals, Meire Fortunato, and Navdeep Jaitly. 2015. Pointer networks. *Advances in Neural Information Processing Systems*, 28:2692–2700.

David Wadden, Ulme Wennberg, Yi Luan, and Hannaneh Hajishirzi. 2019. Entity, relation, and event extraction with contextualized span representations. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pages 5783–5788. Association for Computational Linguistics.

Chenguang Wang, Xiao Liu, Zui Chen, Haoyun Hong, Jie Tang, and Dawn Song. 2022. Deepstruct: Pretraining of language models for structure prediction. In *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 803–823. Association for Computational Linguistics.

Chunqi Wang, Ji Zhang, and Haiqing Chen. 2018. Semi-autoregressive neural machine translation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 479–488. Association for Computational Linguistics.

Ruize Wang, Duyu Tang, Nan Duan, Zhongyu Wei, Xuanjing Huang, Jianshu Ji, Guihong Cao, Daxin Jiang, and Ming Zhou. 2021. K-adapter: Infusing knowledge into pre-trained models with adapters. In *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, Online Event, August 1-6, 2021*, volume ACL/IJCNLP 2021 of *Findings of ACL*, pages 1405–1418. Association for Computational Linguistics.

Yucheng Wang, Bowen Yu, Yueyang Zhang, Tingwen Liu, Hongsong Zhu, and Limin Sun. 2020. Tplinker: Single-stage joint extraction of entities and relations through token pair linking. In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, pages 1572–1582. International Committee on Computational Linguistics.

Zhepei Wei, Jianlin Su, Yue Wang, Yuan Tian, and Yi Chang. 2020. A novel cascade binary tagging framework for relational triple extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 1476–1488. Association for Computational Linguistics.

Xin Xie, Ningyu Zhang, Zhoubo Li, Shumin Deng, Hui Chen, Feiyu Xiong, Mosha Chen, and Huajun Chen. 2022. From discrimination to generation: Knowledge graph completion with generative transformer. In *Companion of The Web Conference 2022, Virtual Event / Lyon, France, April 25 - 29, 2022*, pages 162–165. ACM.

Canwen Xu and Julian J. McAuley. 2022. A survey on dynamic neural networks for natural language processing. *CoRR*, abs/2202.07101.

Hang Yan, Junqi Dai, Tuo Ji, Xipeng Qiu, and Zheng Zhang. 2021a. A unified generative framework for aspect-based sentiment analysis. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 2416–2429. Association for Computational Linguistics.

Hang Yan, Tao Gui, Junqi Dai, Qipeng Guo, Zheng Zhang, and Xipeng Qiu. 2021b. A unified generative framework for various NER subtasks. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 5808–5822. Association for Computational Linguistics.

Hongbin Ye, Ningyu Zhang, Shumin Deng, Mosha Chen, Chuanqi Tan, Fei Huang, and Huajun Chen. 2021. Contrastive triple extraction with generative transformer. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 14257–14265. AAAI Press.

Haiyang Yu, Ningyu Zhang, Shumin Deng, Hongbin Ye, Wei Zhang, and Huajun Chen. 2020. Bridging text and knowledge with multi-prototype embedding for few-shot relational triple extraction. In *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, pages 6399–6410. International Committee on Computational Linguistics.

Dmitry Zelenko, Chinatsu Aone, and Anthony Richardella. 2003. Kernel methods for relation extraction. *J. Mach. Learn. Res.*, 3:1083–1106.

Daojian Zeng, Kang Liu, Yubo Chen, and Jun Zhao. 2015. Distant supervision for relation extraction via piecewise convolutional neural networks. In *EMNLP*, pages 1753–1762.

Daojian Zeng, Haoran Zhang, and Qianying Liu. 2020. Copymtl: Copy mechanism for joint extraction of entities and relations with multi-task learning. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 9507–9514. AAAI Press.

Xiangrong Zeng, Shizhu He, Daojian Zeng, Kang Liu, Shengping Liu, and Jun Zhao. 2019. Learning the extraction order of multiple relational facts in a sentence with reinforcement learning. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3-7, 2019*, pages 367–377. Association for Computational Linguistics.

Xiangrong Zeng, Daojian Zeng, Shizhu He, Kang Liu, and Jun Zhao. 2018. Extracting relational facts by an end-to-end neural model with copy mechanism. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers*, pages 506–514. Association for Computational Linguistics.

Jing Zhang, Bo Chen, Lingxi Zhang, Xirui Ke, and Haipeng Ding. 2021a. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. *AI Open*, 2:14–35.

Ningyu Zhang, Shumin Deng, Zhen Bi, Haiyang Yu, Jiacheng Yang, Mosha Chen, Fei Huang, Wei Zhang, and Huajun Chen. 2020a. Openue: An open toolkit of universal extraction from text. In *EMNLP (Demos)*, pages 1–8.

Ningyu Zhang, Xin Xu, Liankuan Tao, Haiyang Yu, Hongbin Ye, Shuofei Qiao, Xin Xie, Xiang Chen, Zhoubo Li, Lei Li, et al. 2022a. Deepke: A deep learning based knowledge extraction toolkit for knowledge base population. In *Proceedings of the EMNLP Demonstrations*.

Ningyu Zhang, Hongbin Ye, Shumin Deng, Chuanqi Tan, Mosha Chen, Songfang Huang, Fei Huang, and Huajun Chen. 2021b. Contrastive information extraction with generative transformer. *IEEE ACM Trans. Audio Speech Lang. Process.*, 29:3077–3088.

Ranran Haoran Zhang, Qianying Liu, Aysa Xuemo Fan, Heng Ji, Daojian Zeng, Fei Cheng, Daisuke Kawahara, and Sadao Kurohashi. 2020b. Minimize exposure bias of seq2seq models in joint entity and relation extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020*, volume EMNLP 2020 of *Findings of ACL*, pages 236–246. Association for Computational Linguistics.

Shuai Zhang, Yongliang Shen, Zeqi Tan, Yiquan Wu, and Weiming Lu. 2022b. De-bias for generative extraction in unified NER task. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 808–818. Association for Computational Linguistics.

Suncong Zheng, Feng Wang, Hongyun Bao, Yuexing Hao, Peng Zhou, and Bo Xu. 2017. Joint extraction of entities and relations based on a novel tagging scheme. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers*, pages 1227–1236. Association for Computational Linguistics.

Yucheng Zhou, Tao Shen, Xiubo Geng, Guodong Long, and Daxin Jiang. 2022. Claret: Pre-training a correlation-aware context-to-event transformer for event-centric generation and classification. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2559–2575. Association for Computational Linguistics.

Chenguang Zhu, William Hinthorn, Ruochen Xu, Qingkai Zeng, Michael Zeng, Xuedong Huang, and Meng Jiang. 2020. Boosting factual correctness of abstractive summarization with knowledge graph. *CoRR*, abs/2003.08612.

Xiangru Zhu, Zhixu Li, Xiaodan Wang, Xueyao Jiang, Penglei Sun, Xuwu Wang, Yanghua Xiao, and Nicholas Jing Yuan. 2022. Multi-modal knowledge graph construction and application: A survey. *CoRR*, abs/2202.05786.

A Timeline Analysis

As shown in Table 4, we summarize a number of existing research papers in chronological order in the form of a timeline, which hopefully helps researchers who are new to this topic understand the evolution of the generative KGC paradigms.

TABLE 4 Timeline of generative KGC. The time for each paper is based on its first arXiv version (if it exists) or estimated submission time. Works in red consider copy-based sequence methods; works in blue consider structure-linearized sequence methods; works in green consider label-augmented sequence methods; works in orange consider indice-based sequence methods; works in purple consider blank-based sequence methods.

2018.06.15	CopyRE (Zeng et al., 2018)	2021.09.10	TEMPGEN (Huang et al., 2021)
2019.06.12	COMET (Bosselut et al., 2019)	2021.11.07	REBEL (Cabot and Navigli, 2021)
2019.07.28	Seq2Seq4ATE (Ma et al., 2019)	2022.01.17	SQUIRE (Bai et al., 2022)
2019.08.19	Nested-seq (Straková et al., 2019)	2022.02.04	GenKGC (Xie et al., 2022)
2019.11.04	CopyRRL (Zeng et al., 2019)	2022.02.27	EPGEL (Lai et al., 2022)
2019.11.22	PNDec (Nayak and Ng, 2020)	2022.04.11	HuSe-Gen (Saha et al., 2022)
2019.11.24	CopyMTL (Zeng et al., 2020)	2022.05.04	ClarET (Zhou et al., 2022)
2020.01.30	SEQ2SEQ-PTR (Rongali et al., 2020)	2022.05.12	GTEE (Liu et al., 2022)
2020.09.14	CGT (Zhang et al., 2021b)	2022.05.15	X-GEAR (Huang et al., 2022)
2020.09.15	ANL (Athiwaratkun et al., 2020a)	2022.05.22	DEEPSTRUCT (Wang et al., 2022)
2020.10.02	GENRE (Cao et al., 2021)	2022.05.22	De-Bias (Zhang et al., 2022b)
2021.01.01	PolicyIE (Ahmad et al., 2021)	2022.05.22	KGT5 (Saxena et al., 2022)
2021.01.14	TANL (Paolini et al., 2021)	2022.05.22	PAIE (Ma et al., 2022)
2021.04.13	BART-Gen (Li et al., 2021b)	2022.05.23	UIE (Lu et al., 2022)
2021.04.21	GRIT (Du et al., 2021a)	2022.09.15	Seq2rel (Giorgi et al., 2022)
2021.06.02	UGF for NER (Yan et al., 2021b)	2022.09.15	KG-S2S (Chen et al., 2022a)
2021.06.08	UGF for ABSA (Yan et al., 2021a)		
2021.06.11	GTT (Du et al., 2021b)		
2021.06.17	Text2Event (Lu et al., 2021)		
2021.06.30	HySPA (Ren et al., 2021)		
2021.08.29	DEGREE (Hsu et al., 2022)		