

Developer-Intent Driven Code Comment Generation

开发者意图驱动的代码注释生成

沐方文, 陈啸, 石琳, 王松, 王青

The 45th International Conference on Software Engineering (ICSE 2023)

联系人: 沐方文 邮箱: fangwen2020@iscas.ac.cn

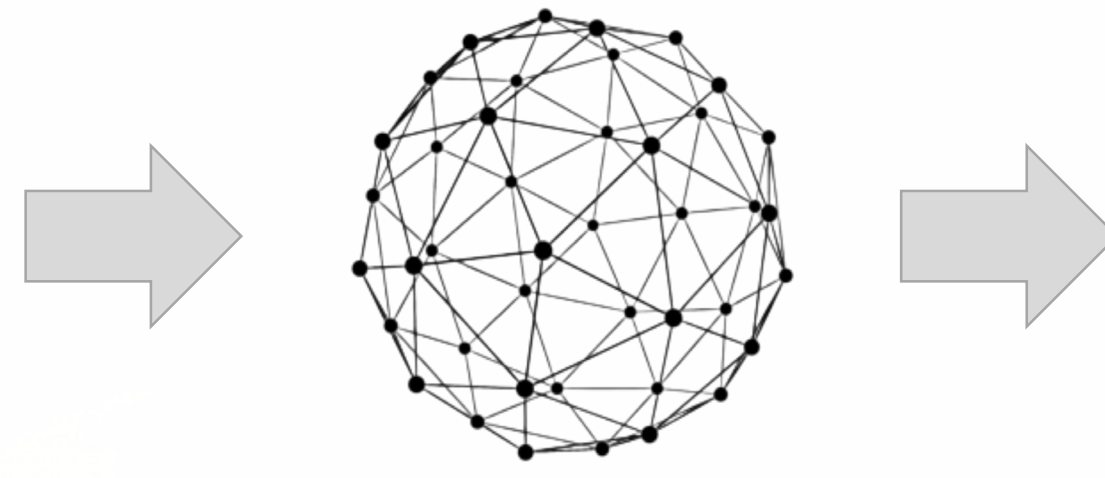
Motivation

Existing Code Comment Generation

No consideration of developers' intent!

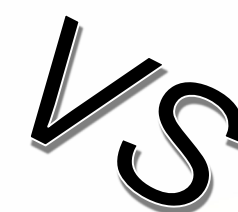
```
public synchronized boolean start() {
    if (!isStarted()) {
        final ExecutorService tempExec;
        executor = getExternalExecutor();
        if (executor == null) {
            executor = tempExec = createExecutor();
        } else {
            tempExec = null;
        }
        future = executor.submit(createTask(tempExec));
        return true;
    }
    return false;
}
```

A Java method *start()*



Rencos (Zhang et al. ICSE20)

starts the execution of the executor
Generated comment

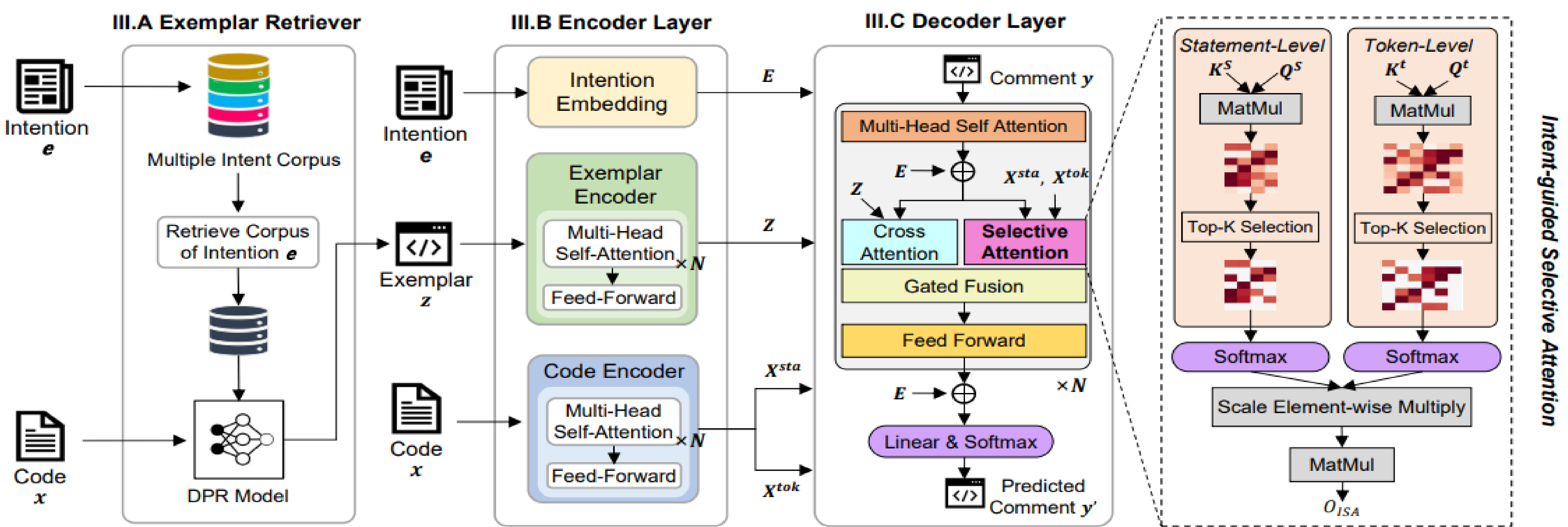


What Developers Want:

Category	Description
What	Describes the functionality of a method
Why	Explains the reason why a method is provided or the design rationale of the method
How-to-use	Describes the usage or the expected set-up of using a method
How-it-is-done	Describes the implementation details of a method
Property	Asserts properties of a method including pre- or post-conditions of a method

Table: The intent taxonomy of code comments (Zhai et al. ICSE 20)

Approach: DOME



- Goal: DOME aims to generate **intent-aware comments** for one code snippet under different intents
- Input: A code snippet & **An intent category**
- Output: The comment that **describes the code based on the given intent**
- Components:
 - **Exemplar Retriever**, for retrieving the most similar comment as the exemplar.
 - **Encoder Layer**, for encoding the code, retrieved exemplar, and intent into vectors.
 - **Decoder Layer**, for leveraging the attention guided by the intent to extract the most relevant information from the code and generating the intent-aware comments.

Evaluation

Intent	Method		Funcom			TLC		
			BLEU	ROUGE-L	METEOR	BLEU	ROUGE-L	METEOR
What	Baseline	Rencos	26.19	31.10	14.61	23.28	36.89	16.02
		EditSum	27.58	31.06	14.24	21.34	33.73	13.42
		AST-Trans	27.84	38.28	18.49	23.42	34.24	17.17
	Our Approach	DOME	33.29	41.67	20.53	25.39	39.56	18.22
		DOME w/o ISA	32.01	38.74	18.65	24.37	37.93	16.51
		DOME w/o ER	31.33	38.12	18.57	23.82	37.24	16.37
Why	Baseline	Rencos	24.81	30.12	14.21	20.55	32.99	14.54
		EditSum	27.17	30.98	14.66	18.42	29.85	11.81
		AST-Trans	25.96	35.74	17.77	19.31	29.37	14.94
	Our Approach	DOME	33.07	42.31	20.56	21.97	35.31	15.77
		DOME w/o ISA	31.79	39.37	19.11	21.41	34.27	15.56
		DOME w/o ER	31.13	38.78	18.99	19.60	32.16	12.71
Usage	Baseline	Rencos	25.54	27.74	13.34	14.62	22.61	10.21
		EditSum	25.25	27.25	12.79	14.00	21.51	9.07
		AST-Trans	24.93	30.90	15.09	13.24	18.20	9.16
	Our Approach	DOME	31.63	39.31	19.34	17.16	26.11	12.36
		DOME w/o ISA	30.57	37.24	17.54	16.74	25.59	11.25
		DOME w/o ER	30.42	37.15	17.37	15.33	24.96	10.94
Done	Baseline	Rencos	19.84	29.28	12.53	18.58	33.73	13.12
		EditSum	22.22	29.73	12.62	16.84	31.18	11.12
		AST-Trans	19.65	33.60	14.40	17.61	30.32	13.29
	Our Approach	DOME	26.98	39.52	17.65	20.48	36.66	14.73
		DOME w/o ISA	26.03	38.19	18.20	19.50	36.57	13.10
		DOME w/o ER	25.78	37.73	18.10	19.14	35.72	13.01
Property	Baseline	Rencos	25.57	35.60	16.39	23.82	38.85	17.76
		EditSum	26.79	36.09	16.60	22.35	37.74	16.33
		AST-Trans	28.29	43.54	20.73	23.54	36.56	18.47
	Our Approach	DOME	34.18	49.43	24.32	26.01	45.73	21.29
		DOME w/o ISA	32.21	46.92	22.70	25.35	43.68	19.47
		DOME w/o ER	31.80	46.09	21.59	25.01	42.85	19.16

Two Benchmarks:

➢ Funcom, TLC

Three Metrics:

➢ BLEU, ROUGE, METEOR

Three SOTA Baselines:

➢ Recos, EditSum, AST-Trans

Results:

➢ Compared to the best baselines, DOME improves the performance of BLEU, ROUGE-L, and METEOR by **25.66%**, **16.59%**, and **18.38%** on Funcom dataset, by **10.06%**, **11.09%**, and **14.93%** on TLC dataset, respectively.

Conclusion

- We propose DOME, which can generate diverse comments given different intents. To our best knowledge, this is **the first work that incorporates developer intents in comment generation**.
- We conduct an experimental evaluation of the performance of DOME, which shows that **DOME outperforms all baselines in main evaluation metrics**.

<https://github.com/ICSE-DOME/DOME>