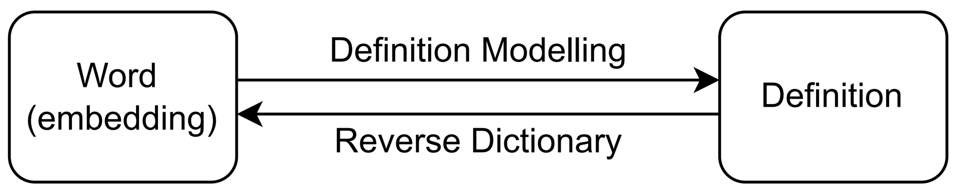
A Unified Model for Reverse Dictionary and Definition Modelling

Pinzhen Chen Zheng Zhao

School of Informatics, University of Edinburgh

{pinzhen.chen,zheng.zhao}@ed.ac.uk

Main Contributions



We present a unified model for two reciprocal tasks: reverse dictionary and definition modelling. The model achieves strong automatic and human evaluation results without relying on external human-annotated data.

The Unified Model

The model learns to encode definitions and words using a shared layer, and then generates both forms via multi-tasking to accomplish reverse dictionary and definition modelling separately. Such a trained system resembles a dual-way neural dictionary.

Unification enabled 1) extra learning objectives like reconstruction and embedding similarity; 2) shared encoder and decoder embeddings.

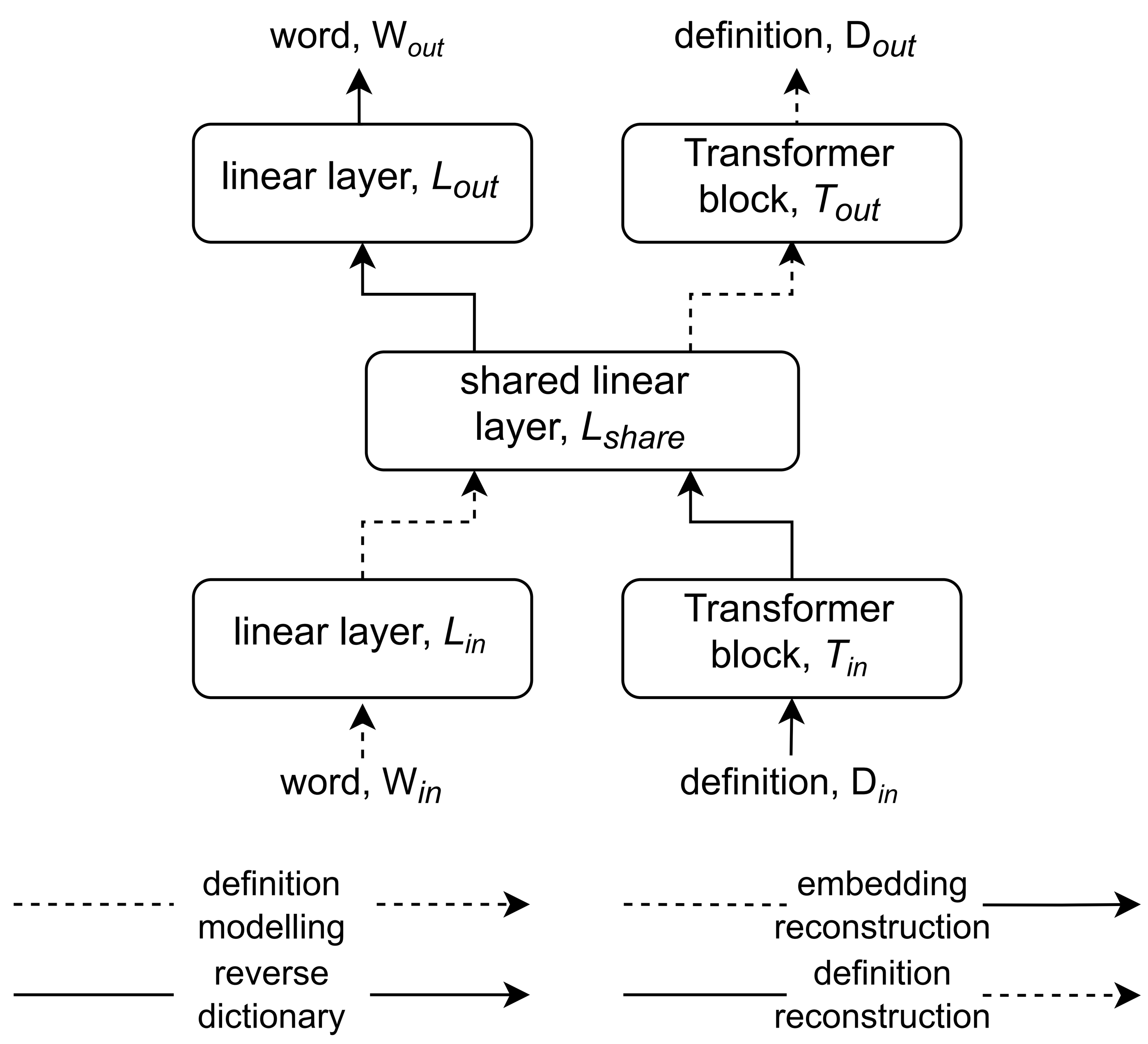


Figure 1. The unified model architecture.

Definition Human Evaluation

Our model notably outperforms a Transformer baseline in both types of human evaluation on definition generation:

1. reference-less: pick the preferred output based on the query word.
2. reference-based: pick the preferred output based on the reference.

	reference-less	reference-based
Transformer	25 (31%)	32 (40%)
unified	**50 (63%)**	**42 (53%)**

Table 1. Chances a model's output is preferred by human evaluators.

Ablation Training Dynamics

We study training objective ablation with the unified model. 1-task refers to using a single reverse dictionary or definition modelling objective; 3-task refers to disabling reconstruction tasks; 5-task is using all objectives.

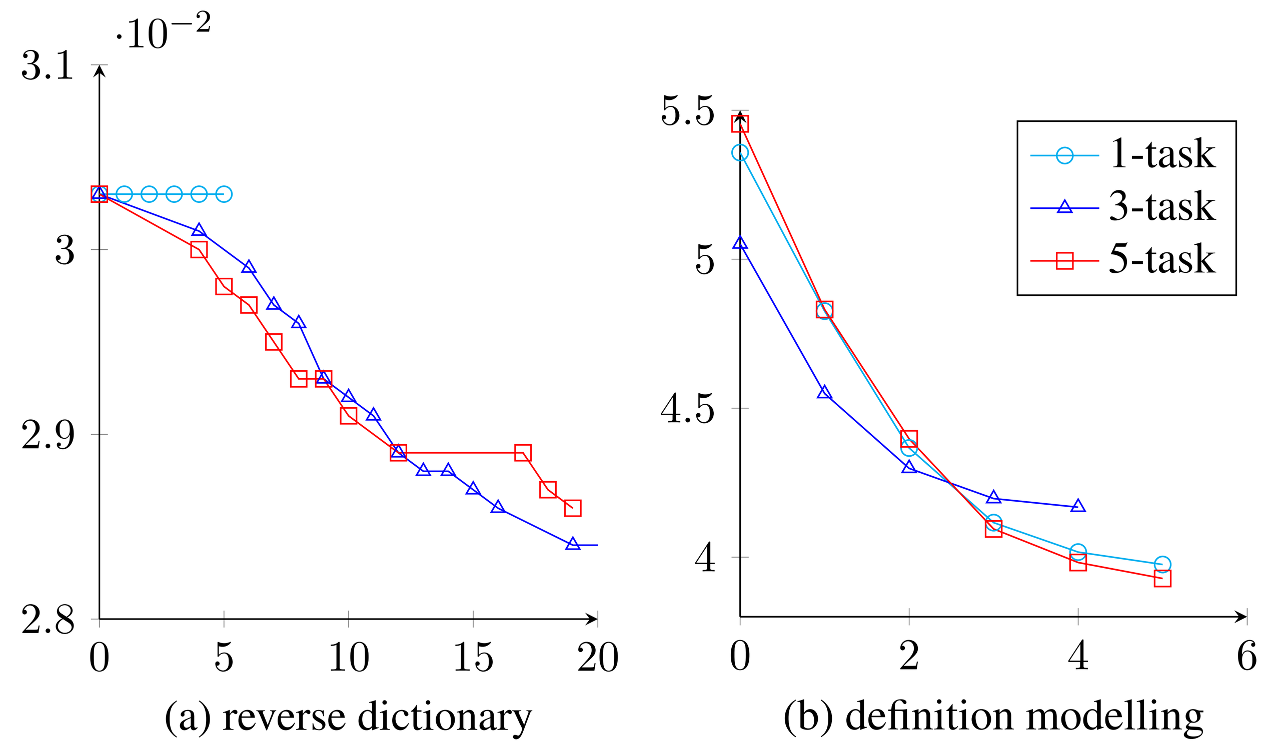


Figure 2. Training losses of the unified model with different objectives

Experiments and Results

	unseen test			human description		
	median rank	acc@ 1/10/100	rank std.	median rank	acc@ 1/10/100	rank std.
OneLook.com	-	-	-	5.5	.33/.54/.76	332
bag-of-words	248	.03/.13/.39	424	22	.13/.41/.69	308
RNN	171	.03/.15/.42	404	17	.14/.40/.73	274
category inference	170	.05/.19/.43	420	16	.14/.41/.74	306
multi-sense	276	.03/.14/.37	426	1000	.01/.04/.18	404
super-sense	465	.02/.11/.31	454	115	.03/.15/.47	396
multi-channel	54	.09/.29/.58	**358**	**2**	**.32**/.64/.88	203
Transformer	79	.01/.14/.59	473	27	.05/.23/.87	332
our unified	**18**	**.13/.39/.81**	386	4	.22/.64/**.97**	**183**
+ share embed	20	.08/.36/.77	410	4	.23/**.65**/**.97**	**183**

(a) **Reverse dictionary** results on Hill et al.'s data with past results from Zhang et al.

	unseen test	
	BLEU	Rouge-L
RNN	1.7	15.8
xSense	2.0	15.9
Transformer	2.4	17.9
our unified	2.2	18.5
+ share embed	**3.0**	**20.2**

(b) **Definition modelling** results on Chang et al.'s data, with past numbers from Chang & Chen's replicate.

Table 2. Experimental results on reverse dictionary (left) and definition modelling (right).

References

Hill et al., Learning to Understand Phrases by Embedding the Dictionary, TACL 2016
Chang et al., xSense: Learning Sense-Separated Sparse Representations and Textual Definitions for Explainable Word Sense Networks, arXiv 2019
Chang & Chen, What Does This Word Mean? Explaining Contextualized Embeddings with Natural Language Definition, EMNLP 2019
Zhang et al., Multi-channel Reverse Dictionary Model, AAAI 2020

Acknowledgement

Pinzhen Chen is funded by the High Performance Language Technologies with Innovate UK. Zheng Zhao is supported by the UKRI Centre for Doctoral Training in Natural Language Processing (EP/S022481/1).

THE UNIVERSITY of EDINBURGH informatics

Innovate UK

NLP UKRI CENTRE FOR DOCTORAL TRAINING