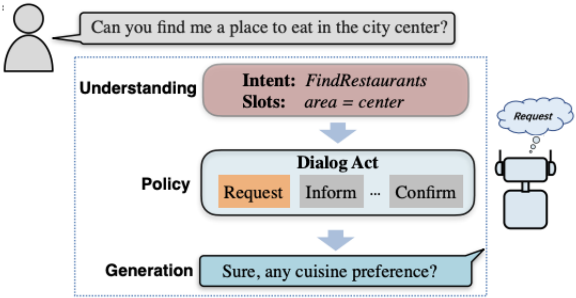
[Review] GALAXY: A Generative Pre-trained Model for Task-Oriented Dialog with Semi-supervised Learning and Explicit Policy Injection

Vishal Anand

Columbia University, Microsoft Corporation

September 28, 2022

Motivation



- Task-Oriented Dialog (TOD) Systems
- Construct Mechanism ⇒
  1. Dialog Understanding: *structured-semantics* from user utterances
  2. Policy Planning: Dialog Act (*DA*) that leads to task success
  3. Dialog Generation: Produce responses

Contribution

1. First semi-supervised study on pre-training to model explicit dialog policy for PCMs (Pre-trained Conversation Models)
2. GALAXY learns dialog policy knowledge $\Rightarrow$ SOTA on TOD
3. New Labeled dataset *UniDA* and unlabeled dialog corpus *UnDial*

TL;DR - Takeaway

Optimization Equations

$$\mathcal{L}_{pre} = \mathcal{L}_{unlabel} + \mathcal{L}_{label} \tag{1}$$

$$\mathcal{L}_{label} = \mathcal{L}_{RS} + \mathcal{L}_{RG} + \mathcal{L}_{DA} + \mathcal{L}_{KL} \tag{2}$$

$$\mathcal{L}_{unlabel} = \mathcal{L}_{RS} + \mathcal{L}_{RG} + g\mathcal{L}_{KL} \tag{3}$$

$$\mathcal{L}_{fine} = \mathcal{L}_{RS} + \mathcal{L}_{RG} + \alpha\mathcal{L}_{DA} \tag{4}$$

- *pre*: Pre-training objectives
  - *RS*: Response Selection
  - *RG*: Response Generation
  - *DA*: Dialog Act Prediction
  - *KL*: Consistency Regularization (Kullback-Leibler)
- *fine*: Semi-supervised learning from labeled+unlabeled data

Related Work

- Pre-trained Language Models (PLMs): Large Scale text corpora (Transformers) e.g. SimpleTOD
- Pre-trained Conversation Models (PCMs): Variants of PLMs:
  - Train on Dialog Data
  - Designed on new dialog-oriented pre-training objectives
  - Integrate both of the former points together
- Semi-supervised Learning (SSL)
  - Labeled and Unlabeled data
  - Early results from generative models - VAEs, GANs
  - Pseudo-Labeling: Further training on unlabeled data after initial tagging via label-data-trained-model.
    (Shared lower layers, task-specific top-layers)
  - Consistency Regularization: Perturb unlabeled data-points, and minimize KL-divergence.

Input Representation

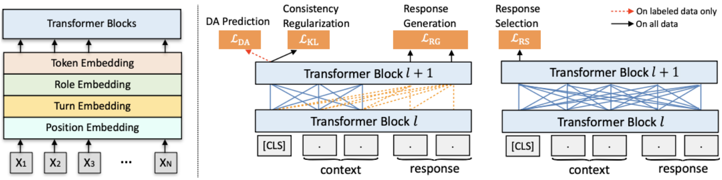


Figure 1: Architecture of our pre-trained dialog model. The left part illustrate the input representations, which contain embeddings of tokens, roles, turns, and positions. The right part shows the pre-trained objectives. Blue lines denote the bi-directional attention. Dashed yellow lines denote the uni-directional attention.

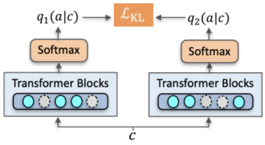


Figure 2: The procedure of computing L_{KL}.

Source: He et al. GALAXY paper.

Pre-training objectives - Response Selection

- Binary Classification problem: Given a (context, response) pair from corpus, perform negative sampling.
- sigmoid(bidirectional_encoder(c,r))

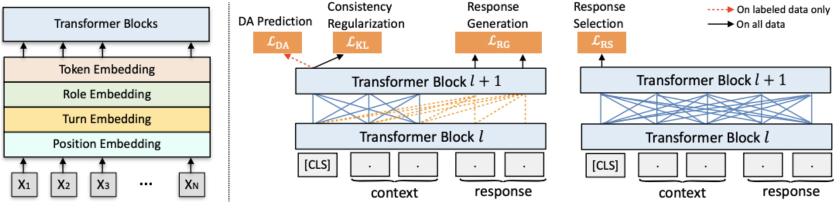


Figure 3: The procedure of computing L_{RS}.

Source: He et al. GALAXY paper.

Pre-training objectives - Response Generation

- Auto-regressive (based on past values) dialog response prediction on dialog context c.
- To minimize negative Log likelihood loss

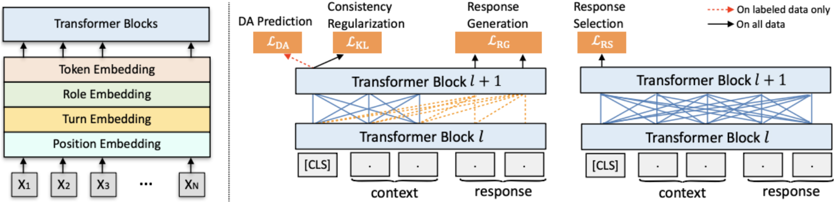


Figure 4: The procedure of computing L_{RG}.

Source: He et al. GALAXY paper.

Pre-training objectives - Dialog Act Prediction

- Predict DA label a from context c (multi-mapping).
- Encode the objective with sigmoid of a Bernoulli distribution over labels given context, since $\Sigma p(a|c)$ can be more than one.

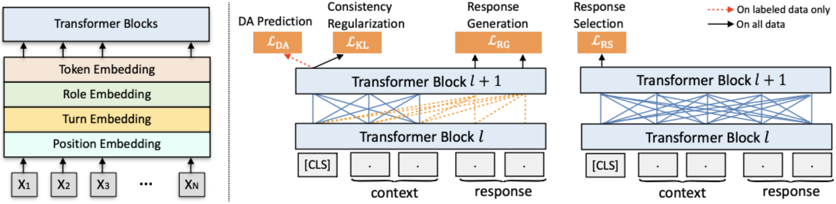


Figure 5: The procedure of computing L_{DA}.

Source: He et al. GALAXY paper.

Pre-training objectives - Consistency Regularization

- Labels are absent by design in UnDial
- Employ dropouts on given context twice on feed-forward network, and minimize bidirectional KL-divergence of distribution

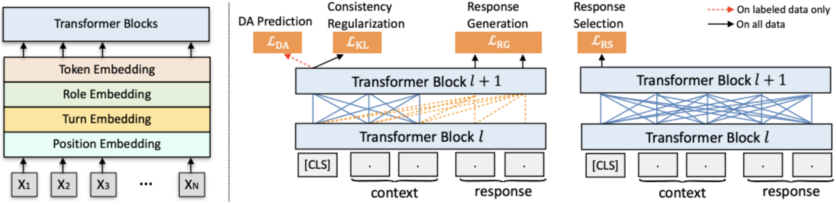


Figure 6: The procedure of computing L_{KL}.

Source: He et al. GALAXY paper.

Semi-Supervised Pre-trained mechanism

$$\mathcal{L}_{label} = \mathcal{L}_{RS} + \mathcal{L}_{RG} + \mathcal{L}_{DA} + \mathcal{L}_{KL} \tag{5}$$

$$\mathcal{L}_{unlabel} = \mathcal{L}_{RS} + \mathcal{L}_{RG} + g\mathcal{L}_{KL} \tag{6}$$

$$\mathcal{L}_{pre} = \mathcal{L}_{unlabel} + \mathcal{L}_{label} \tag{7}$$

- Discard noisy samples from unlabeled data (examples follow in a later slide)
- Use perturbed distribution as proxy for current entropy.

$$g = \min\left\{\max\left\{0, \frac{E_{max} - (E + \log E)}{E_{max}}\right\}, 1\right\}$$

Semi-Supervised Fine-tuning

$$\mathcal{L}_{fine} = \mathcal{L}_{RS} + \mathcal{L}_{RG} + \alpha\mathcal{L}_{DA} \tag{8}$$

- α is one when we have annotation, else zero
- r* = (d, r), where d comprises of belief states and dialog acts.

Model Architecture Reference - UniLM

Dong et al. Unified Language Model Pre-training for Natural Language Understanding and Generation. NeurIPS 2019.

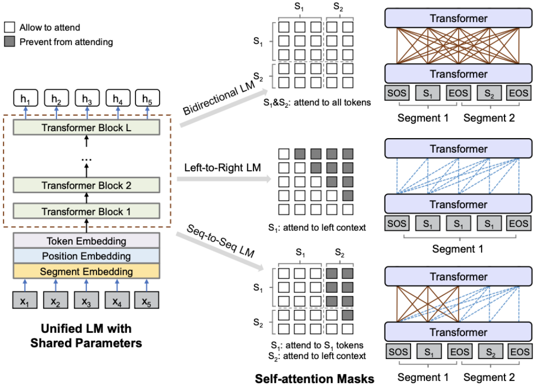


Figure 1: Overview of unified LM pre-training. The model parameters are shared across the LM objectives (i.e., bidirectional LM, unidirectional LM, and sequence-to-sequence LM). We use different self-attention masks to control the access to context for each word token. The right-to-left LM is similar to the left-to-right one, which is omitted in the figure for brevity.

Source: Dong et al. UniLM paper.

Model Architecture Reference - UniLM

Dong et al. Unified Language Model Pre-training for Natural Language Understanding and Generation. NeurIPS 2019.

- Bi-directional encoder for understanding
- Uni-directional decoder for generation
- Encoder and Decoder share weights

Input Representation Reference - PLATO

Dong et al. PLATO: Pre-trained Dialogue Generation Model with Discrete Latent Variable. NAACL 2020.

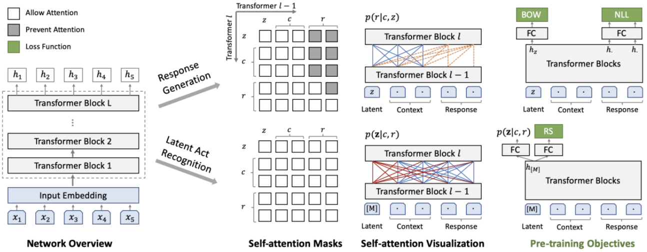


Figure 2: Architecture of dialogue generation with discrete latent variable. In self-attention visualization, red and blue lines denote bi-directional attention, and dashed orange lines denote uni-directional attention.

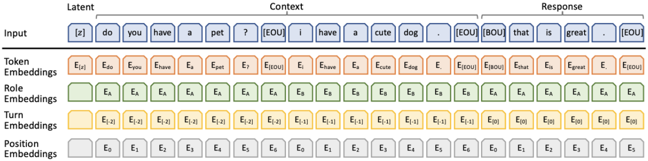


Figure 3: Input representation. The input embedding is the sum of token, role, turn and position embeddings.

Source: Bao et al. PLATO paper.

Pre-Training Dialog Labeled Dataset: *UniDA*

Name	# Dialogs	# Utterance	# Unified DA
MultiWOZ	10,433	142,968	11
Frames	1,369	19,986	14
MSRe2e	10,087	74,686	12
SGD	22,825	463,284	9
DSTC2	3,235	44,532	7
SimJoint	3,008	24,112	6
STAR	6,652	107,846	11
DailyDialog	13,117	98,366	9
UniDA	70,726	975,780	20
Unified DAs	*request, select, reqalts, affirm, not_sure, inform, impl-confirm, expl-confirm, notify_success, notify_failure, hi, bye, negate, repeat, welcome, thank_you, direct, dont_understand, propose, offer*		

- N.B. Dialog Acts are treated as the annotations disregarding contents
- Text-only TOD systems considered
- Prior universal task-datasets omit features: *not-sure*, *dont-understand*
- 20 frequently used DAs following Bunt et al. ISO. LREC 2010 *

Table Source: He et al. GALAXY paper.

Pre-Training Dialog Unlabeled Dataset: *UnDial*

# Datasets	14
# Dialog Sessions	14M
# Utterances	35M
Avg. Utterances per Dialog	2.5
Avg. Tokens per Utterance	14.6

- Source: Online forums, chat logs, customer service conversations
- Processing on 14 available datasets to generate the said dataset

Table Source: He et al. GALAXY paper.

Discussion

1. How does the semi-supervised method work during the pre-training process?
2. How much improvements does $\mathcal{L}_{DA}$, $\mathcal{L}_{KL}$ and the gating mechanism contribute?
3. How can the model improve task completion in real cases?

Metrics tracked

1. *Inform rate*: How often the entities provided by the system are correct (proxy of precision*)
2. *Success* rate: Rate of system answering all the requested attributes
3. *BLEU score*: fluency of generated responses
4. *Comb score*: (Inform + Success)*0.5 + BLEU
5. *Comb score**: (SuccessF1 + Match)*0.5 + BLEU
   [In-Car: Lei et al. (2018)]

Evaluation - Benchmarks

Model	MultiWOZ2.0				MultiWOZ2.1			
	Inform	Success	BLEU	Comb	Inform	Success	BLEU	Comb
SimpleTOD (Hosseini-Asl et al. 2020)	84.40	70.10	15.01	92.26	85.00	70.50	15.23	92.98
DoTS (Jeon and Lee 2021)	86.59	74.14	15.06	95.43	86.65	74.18	15.90	96.32
SOLOIST (Peng et al. 2020a)	85.50	72.90	16.54	95.74	–	–	–	–
MinTL (Lin et al. 2020)	84.88	74.91	17.89	97.79	–	–	–	–
PPTOD (Su et al. 2021)	89.20	79.40	18.62	102.92	87.09	79.08	19.17	102.26
UBAR (Yang, Li, and Quan 2020)	**95.40**	80.70	17.00	105.05	**95.70**	81.80	16.50	105.25
GALAXY(w/o pre-train)	93.10	81.00	18.44	105.49	93.50	81.70	18.32	105.92
GALAXY	94.40	**85.30**	**20.50**	**110.35**	95.30	**86.20**	**20.01**	**110.76**

Figure 7: performances on MultiWOZ2.0/2.1.

Model	Match	SuccF1	BLEU
SEDST (Jin et al. 2018)	84.50	82.90	19.30
TSCP (Lei et al. 2018)	84.50	81.10	21.90
LABES (Zhang et al. 2020a)	**85.80**	77.00	22.80
FSDM (Shu et al. 2019)	84.80	82.10	21.50
GALAXY (w/o pre-train)	81.90	83.30	22.00
GALAXY	85.30	**83.60**	**23.00**

Figure 8: performances on In-Car.

Evaluation - Benchmarks

Model	Inform	Success	BLEU	Comb
UniLM	92.40	81.40	18.45	105.35
PLATO	91.20	77.20	16.68	100.88
TOD-UniLM	93.50	81.30	19.13	106.53
TOD-PLATO	92.10	79.40	17.23	102.98
GALAXY	**94.40**	**85.30**	**20.50**	**110.35**

Figure 10: performances of different pre-trained conversation models on MultiWOZ2.0.

Image Source: He et al. GALAXY paper.

Semi-Supervised Evaluation: Low Resource Analysis

Model	5% data			10% data			20% data		
	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU
DAMD	56.60	24.50	10.60	62.00	39.40	14.50	77.90	70.30	12.10
SOLOIST	69.30	52.30	11.80	69.90	51.90	14.60	74.00	60.10	15.24
MinTL	75.48	60.96	13.98	78.08	66.87	15.46	82.48	68.57	13.00
PPTOD	79.86	63.48	14.89	84.42	68.36	15.57	84.94	71.70	17.01
UBAR	73.04*	60.28*	16.03*	79.20*	68.70*	16.09*	82.50*	66.60*	**17.72***
GALAXY	**80.59**	**67.43**	**17.39**	**87.00**	**75.00**	**17.65**	**89.55**	**75.85**	17.54

End to end results of low-resource experiments. 5% (400 dialogs), 10% (800 dialogs), 20% (1600 dialogs) of training data is used to train each model.

- More sample efficient than recent pre-trained models.

Table Source: He et al. GALAXY paper.

Semi-Supervised Evaluation: Pre-Trained Methods

- GALAXY$_{multi}$ discards $\mathcal{L}_{KL}$.

Model	Inform	Success	BLEU	Comb
Pseudo-Labeling	90.10	80.30	16.79	101.99
VAE	89.00	76.40	16.48	99.18
GALAXY$_{multi}$	93.90	82.30	19.17	107.27
GALAXY	**94.40**	**85.30**	**20.50**	**110.35**

Table 6: E2E performance of different semi-supervised pre-training methods on MultiWOZ2.0.

Table Source: He et al. GALAXY paper.

Discussion - Learning Curve

1. 10% of UniDA and 100% of UnDial used for training, rest is testing set
2. Given limited data, model collapses, outputs become constant.
3. $\mathcal{L}_{DA}$ handles the problem in next iterations.
4. $\mathcal{L}_{KL}$ avoids overfitting.

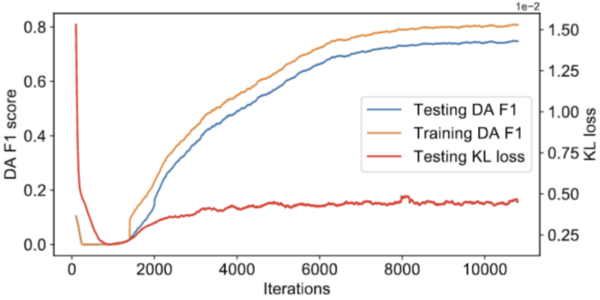


Figure 11: Learning curves of train/test DA F1 scores and the $\mathcal{L}_{KL}$ loss.

Image Source: He et al. GALAXY paper.

Discussion - Ablation Results

Model	Inform	Success	BLEU
GALAXY	**94.40**	**85.30**	**20.50**
$-g$	94.20	83.50	19.26
$-\mathcal{L}_{\mathrm{DA}}$	89.10	79.90	18.77
$-\mathcal{L}_{\mathrm{KL}}$	93.90	82.30	19.17
$-\mathcal{L}_{\mathrm{DA}} - \mathcal{L}_{\mathrm{KL}}$	93.30	81.20	19.54

Figure 12: Ablation results on MultiWoz2.0

- Minus $\mathcal{L}_{DA} \Rightarrow$ collapsing problem occurs
- Minus $\mathcal{L}_{KL} \Rightarrow$ Multitask learning (inadequate unlabeled data use*)
- Minus $\mathcal{L}_{KL}$, $\mathcal{L}_{DA} \Rightarrow$ Pre-training on high quality dialog data

Context	i need either the email address , or just zip code. **(Gate: 1.0)**
Response	zip code : 24627. **(DA: inform)**
Context	i need to return an item , can you help me? **(Gate: 0.91)**
Response	sure , may i have your name please? **(DA: request)**
Context	i pour a little liquor out for habeas. **(Gate: 0.41)**
Response	i pour it into corpus. **(DA: N.A.)**
Context	one word : justice. **(Gate: 0.19)**
Response	let me guess , you drive a 1980 ford pinto. **(DA: N.A.)**

Figure 13: Gating mechanism (g) in effect on unlabeled data.

Image and Table Source: He et al. GALAXY paper.

System Comparative conversation example

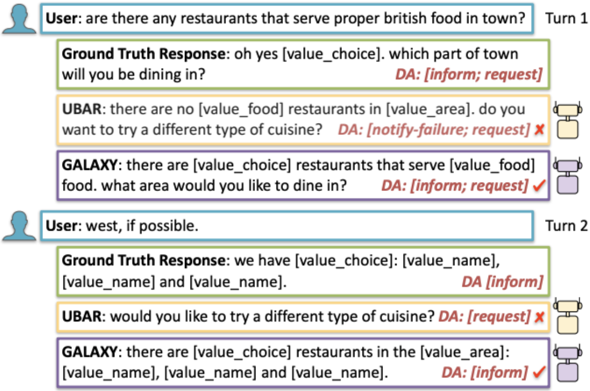


Figure 14: Delexicalized responses by GALAXY and UBAR on MultiWOZ-2.0

Image Source: He et al. GALAXY paper.