Improving Generative Visual Dialog by Answering Diverse Questions

Vishvak Murahari[1] **Prithvijit Chattopadhyay**[1]
Dhruv Batra[1,2] **Devi Parikh**[1,2] **Abhishek Das**[1]

[1]Georgia Tech [2]Facebook AI Research

Abstract

Prior work on training generative Visual Dialog models with reinforcement learning (Das et al., 2017b) has explored a Q-BOT-A-BOT image-guessing game and shown that this 'self-talk' approach can lead to improved performance at the downstream dialog-conditioned image-guessing task. However, this improvement saturates and starts degrading after a few rounds of interaction, and does not lead to a better Visual Dialog model. We find that this is due in part to repeated interactions between Q-BOT and A-BOT during self-talk, which are not informative with respect to the image. To improve this, we devise a simple auxiliary objective that incentivizes Q-BOT to ask diverse questions, thus reducing repetitions and in turn enabling A-BOT to explore a larger state space during RL *i.e.* be exposed to more visual concepts to talk about, and varied questions to answer. We evaluate our approach via a host of automatic metrics and human studies, and demonstrate that it leads to better dialog, *i.e.* dialog that is more diverse (*i.e.* less repetitive), consistent (*i.e.* has fewer conflicting exchanges), fluent (*i.e.* more human-like), and detailed, while still being comparably image-relevant as prior work and ablations.

1 Introduction

Our goal is to build agents that can *see* and *talk i.e.* agents that can perceive the visual world and communicate this understanding in natural language conversations in English. To this end, Das et al. (2017a); de Vries et al. (2017) proposed the task of Visual Dialog – given an image, dialog history consisting of a sequence of question-answer pairs, and a follow-up question about the image, predict a free-form natural language answer to the question – along with a dataset and evaluation metrics[1].

Posing Visual Dialog as a supervised learning problem is unnatural. This is because at every round of dialog, the agent's answer prediction is thrown away and it gets access to ground-truth dialog history, thus disabling it from steering conversations during training. This leads to compounding errors over long-range sequences at test time, a problem also common in training recurrent neural networks for language modeling (Bengio et al., 2015; Ross et al., 2011; Ranzato et al., 2016).

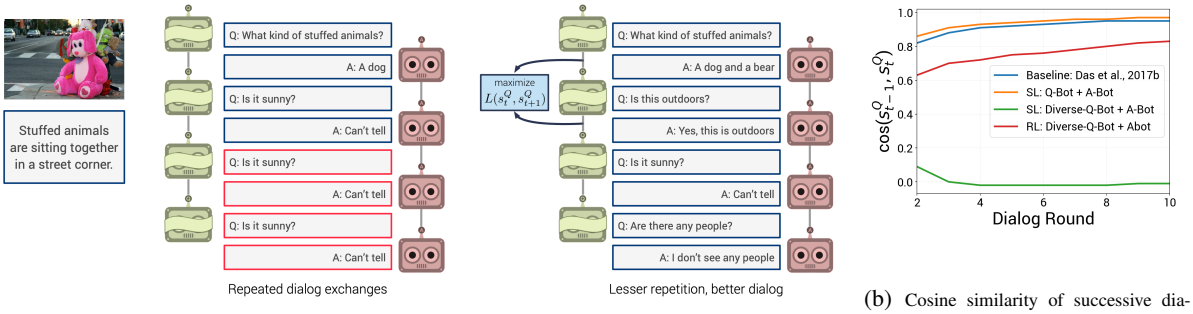
To overcome this, Das et al. (2017b) devised a goal-driven approach for training Visual Dialog agents with deep reinforcement learning. This is formulated as a game between two agents – Q-BOT (that asks questions) and A-BOT (that answers questions). Q-BOT is shown a one-line description of an unseen image that A-BOT has access to and Q-BOT is allowed to ask questions (in natural language) to A-BOT for a fixed number of rounds and simultaneously make predictions of the unseen image. Both agents are rewarded for Q-BOT's image-guessing performance and trained with REINFORCE (Williams, 1992) to optimize this reward. Thus, there is incentive for Q-BOT to ask questions informative of the hidden image, and for A-BOT to provide meaningful answers.

While this reinforcement learning approach leads to improved performance on the image-guessing task than supervised learned agents, it has a few shortcomings – 1) image-guessing performance degrades after a few rounds of dialog (Fig. 2e), and 2) these improvements over supervised learning do not translate to an improved A-BOT, *i.e.* responses from this Visual Dialog agent are not necessarily better (on automatic metrics or human judgements), just that the Q-BOT-A-BOT pair is sufficiently in sync to do well at image-guessing.

We begin by understanding why this is the case, and find that Q-BOT-A-BOT dialog during 'self-talk' often tends to be repetitive *i.e.* the same question-answer pairs get repeated across rounds (Fig. 1a left). Since repeated interactions convey no additional information, image-guessing perfor-

[1] visualdialog.org, guesswhat.ai

Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, pages 1449–1454, Hong Kong, China, November 3–7, 2019.



(a) **Left**. Prior work on training generative Visual Dialog models with RL on an image-guessing task between Q-BOT and A-BOT (Das et al., 2017b) leads to repetitive dialog. **Right**. We devise an auxiliary objective that incentivizes Q-BOT to ask diverse questions, thus reducing repetitions and enabling A-BOT to be exposed to more varied questions during RL, overall leading to better dialog, as measured by automatic metrics and human studies.

(b) Cosine similarity of successive dialog state embeddings within Q-BOT. Prior work (Das et al., 2017b) has high similarity. Our approach explicitly minimizes this similarity leading to more diverse dialog.

Figure 1

mance saturates, and even starts to degrade as the agent forgets useful context from the past.

These repetitions are due to high similarity in Q-BOT's context vectors of successive rounds driving question generation (Fig. 1b). To address this, we devise a smooth-L1 penalty that penalizes similarity in successive state vectors (Section 3). This incentivizes Q-BOT to ask diverse questions, thus reducing repetitions and in turn enabling A-BOT to explore a larger part of the state space during RL *i.e.* be exposed to more visual concepts to talk about, and varied questions to answer (Section 6).

Note that a trivial failure mode with this penalty is for Q-BOT to start generating diverse but totally image-irrelevant questions, which are not useful for the image-guessing task. A good balance between diversity and image-relevance in Q-BOT's questions is necessary to improve at this task.

We extensively evaluate each component of our approach against prior work and baselines:

- Q-BOT on diversity and image-relevance of generated questions during Q-BOT-A-BOT self-talk. We find that diverse-Q-BOT asks *more novel questions* while still being image-relevant.

- Q-BOT-A-BOT self-talk on consistency, fluency, level of detail, and human-interpretability, through automatic metrics and human studies. We find that diverse-Q-BOT-A-BOT dialog after RL is *more consistent, fluent, and detailed*.

- A-BOT on precision and recall of generated answers on the VisDial dataset (Das et al., 2017a) *i.e.* quality of answers to human questions. Training diverse-Q-BOT + A-BOT with RL does not lead to a drop in accuracy on VisDial.

2 Preliminaries

We operate in the same setting as Das et al. (2017b) – an image-guessing task between a questioner (Q-BOT) and an answerer (A-BOT) – where Q-BOT has to guess the image A-BOT has access to by asking questions in multi-round dialog.

We adopt the same training paradigm[2] consisting of 1) a supervised pre-training stage where Q-BOT and A-BOT are trained with Maximum Likelihood Estimation objectives on the VisDial dataset (Das et al., 2017a), and 2) a self-talk RL finetuning stage where Q-BOT and A-BOT interact with each other and the agents are rewarded for each successive exchange based on incremental improvements in guessing the unseen image. We learn parameterized policies $\pi_{\theta_Q}(q_t|s_{t-1}^Q)$ and $\pi_{\theta_A}(a_t|s_t^A)$ for Q-BOT and A-BOT respectively which decide what tokens to utter (*actions*: question q_t and answer a_t at every dialog round t) conditioned on the context available to the agent (*state* representations: s_{t-1}^Q, s_t^A). Q-BOT additionally makes an image feature prediction $\hat{y}_t$ at every dialog round, and the reward is $r_t = ||y^{gt} - \hat{y}_{t-1}||_2^2 - ||y^{gt} - \hat{y}_t||_2^2$, *i.e.* change in distance to the true representation y^{gt} before and after a dialog round. We use REINFORCE (Williams, 1992) to update agent parameters, *i.e.* Q-BOT and A-BOT are respectively updated with $\mathbb{E}_{\pi_Q,\pi_A}[r_t \nabla_{\theta_Q} \log \pi_Q(q_t|s_{t-1}^Q)]$ and $\mathbb{E}_{\pi_Q,\pi_A}[r_t \nabla_{\theta_A} \log \pi_A(a_t|s_t^A)]$ as gradients.

Our transition from supervised to RL is gradual – we supervise for N rounds and have policy-gradient updates for the remaining $10-N$, starting from $N = 9$ till $N = 4$, one round at a time. After reaching $N = 4$, repeating this procedure from $N = 9$ led to further (marginal) improvements.

[2] github.com/batra-mlp-lab/visdial-rl

Both Q-BOT and A-BOT are modeled as Hierarchical Recurrent Encoder-Decoder architectures (Serban et al., 2016). Q-BOT's fc7 (Simonyan and Zisserman, 2015) feature prediction of the unseen image ($\hat{y}$) is conditioned on the dialog history so far (s^Q_{t-1}) via a regression head[3].

3 Smooth-L1 Penalty on Question Repetition

Our goal is to encourage Q-BOT to ask a diverse set of questions so that when A-BOT is exposed to the same during RL finetuning, it is better able to explore its state space[4]. Furthermore, asking diverse questions allows Q-BOT-A-BOT exchanges across rounds to be more informative of the image, thus more useful for the image-guessing task.

We observe that agents trained using the paradigm proposed by Das et al. (2017b) suffer from repetition of *context* across multiple rounds of dialog – similar dialog state embeddings across multiple rounds leading to repeated utterances and similar predicted image representations, which consequently further increases similarity in state embeddings. Fig. 1b shows increasing $\cos(s^Q_{t-1}, s^Q_t)$ across dialog rounds for Das et al. (2017b).

To encourage Q-BOT to ask diverse questions, we propose a simple auxiliary loss that penalizes similar dialog state embeddings. Specifically, given Q-BOT states – s^Q_{t-1}, s^Q_t – in addition to maximizing likelihood of question (during supervised pretraining), or image-guessing reward (during self-talk RL finetuning), we maximize a smooth-L1 penalty on $\Delta_t = \text{abs}(\left\|s^Q_{t-1}\right\|_2 - \left\|s^Q_t\right\|_2)$,

$$f(\Delta_t) = \begin{cases} 0.5\Delta_t^2 & \text{if } \Delta_t < 0.1 \\ 0.1(\Delta_t - 0.05) & \text{otherwise} \end{cases} \quad (1)$$

resulting in $\sum_{t=2}^{N} f(\Delta_t)$ as an additional term in the overall objective (N = no. of dialog rounds).

Note that in order to maximize this penalty, Q-BOT has to push s^Q_{t-1} and s^Q_t further apart, which can only happen if s^Q_{t-1} is updated using a question-answer pair that is different from the previous exchange, thus overall forcing Q-BOT to ask different questions in successive dialog rounds. Similar diversity objectives have also been explored in Li et al. (2016b) as reward heuristics.

[3]Please refer to appendix for architecture details.

[4]A-BOT's state-space is characterized by a representation of the question, image, and the dialog history so far.

Before arriving at (1), and following Fig. 1b, we also experimented with directly minimizing cosine similarity, $\cos(s^Q_{t-1}, s^Q_t)$. This led to the network learning large biases to flip the direction of successive s^Q_{t-1} vectors (without affecting norms), leading to question repetitions in alternating rounds.

4 Experiments

Baselines and ablations. To understand the effect of the proposed penalty, we compare our full approach – '`RL: Diverse-Q-bot + A-bot`' – with the baseline setup in Das et al. (2017b), as well as several ablations – 1) '`SL: Q-bot + A-bot`': supervised agents (*i.e.* trained on VisDial data under MLE, no RL, no smooth-L1 penalty). Comparing to this quantifies how much our penalty + RL helps. 2) '`SL: Diverse-Q-bot + A-bot`': supervised agents where Q-BOT is trained with the smooth-L1 penalty. This quantifies gains from RL.

Automatic Metrics. To evaluate Q-BOT's *diversity* (Table 1), we generate Q-BOT-A-BOT dialogs (with beam size = 5) for 10 rounds on VisDial v1.0 val and compute 1) Novel Questions: the number of new questions (via string matching) in the generated dialog not seen during training, 2) Unique Questions: no. of unique questions per dialog instance (so $\leqslant 10$), 3) Dist-n and Ent-n (Zhang et al., 2018; Li et al., 2016a): the number and entropy of unique n-grams in the generated questions normalized by the total number of tokens, and 4) Mutual Overlap (Deshpande et al., 2019): BLEU-4 overlap of every question in the generated 10-round dialog with the other 9 questions, followed by averaging these 10 numbers. To measure Q-BOT's *relevance*, we report the negative log-likelihood under the model of human questions from VisDial. We evaluate A-BOT's answers to human questions from the VisDial dataset on the retrieval metrics introduced by Das et al. (2017a) (Table 2). Finally, we also evaluate performance of the Q-BOT-A-BOT pair at image-guessing (Fig. 2e), which is the downstream task they are trained with RL for.

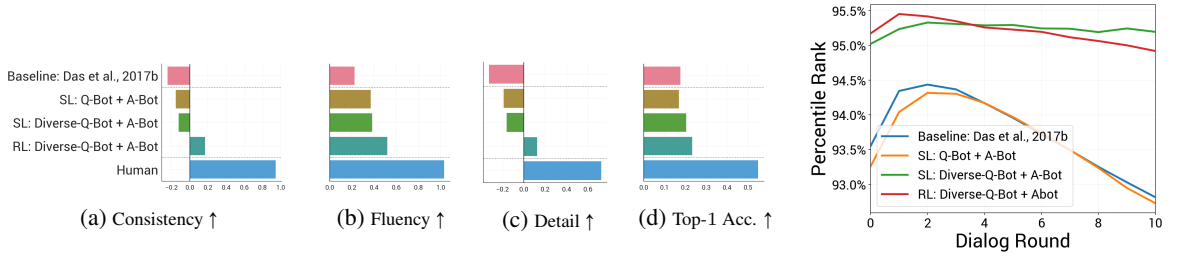
Human Studies. To evaluate how human-understandable Q-BOT-A-BOT dialogs are, we conducted a study where we showed humans these dialogs (from our agents as well as baselines), along with a pool of 16 images from the VisDial v1.0 test-std split – consisting of the unseen image, 5 nearest neighbors (in fc7 space), and 10 random images – and asked humans to pick their top-5 guesses for the unseen image. Our hypothesis

	Diversity							Relevance
	# Novel questions ↑	# Unique questions ↑	Mutual overlap ↓	Ent-1 ↑	Ent-2 ↑	Dist-1 ↑	Dist-2 ↑	Negative log likelihood ↓
Baseline: Das et al. (2017b)	71	6.70 ± 0.07	0.58 ± 0.01	2.72 ± 0.01	3.03 ± 0.02	0.35 ± 0.0	0.43 ± 0.0	**9.94**
SL: Q-BOT + A-BOT	51	6.57 ± 0.07	0.60 ± 0.01	2.70 ± 0.01	3.00 ± 0.02	0.34 ± 0.0	0.42 ± 0.0	10.05
SL: Diverse-Q-BOT + A-BOT	146	7.45 ± 0.07	0.51 ± 0.01	2.82 ± 0.01	3.18 ± 0.01	0.38 ± 0.0	0.48 ± 0.0	10.10
RL: Diverse-Q-BOT + A-BOT	**449**	**8.19** ± 0.06	**0.41** ± 0.01	**2.90** ± 0.01	**3.31** ± 0.01	**0.40** ± 0.0	**0.53** ± 0.0	10.80

Table 1: Q-BOT diversity and relevance on v1.0 val. ↑ indicates higher is better. ↓ indicates lower is better.

	v1.0 val						v1.0 test-std					
	NDCG ↑	MRR ↑	R@1 ↑	R@5 ↑	R@10 ↑	Mean Rank ↓	NDCG ↑	MRR ↑	R@1 ↑	R@5 ↑	R@10 ↑	Mean Rank ↓
Baseline: Das et al. (2017b)	53.76	46.35	36.22	56.15	62.41	19.34	51.60	45.67	35.05	56.30	63.25	19.15
SL: A-BOT	53.10	46.21	36.11	55.82	62.22	19.58	51.18	45.43	34.88	55.65	63.20	19.16
RL: A-BOT (finetuned with Diverse-Q-BOT)	53.91	46.46	36.31	56.26	62.53	19.35	51.67	45.64	34.85	56.55	63.43	18.96

Table 2: A-BOT performance on VisDial v1.0 (Das et al., 2017a). ↑ indicates higher is better. ↓ indicates lower is better.



(a) Consistency ↑ (b) Fluency ↑ (c) Detail ↑ (d) Top-1 Acc. ↑ (e) Q-BOT-A-BOT image-guessing task performance

Figure 2: (**a-d**): Human evaluation of Q-BOT-A-BOT dialog over 50 images and 200 human subjects for each model variant. (**e**): Percentile rank (higher is better) of the true image (shown to A-BOT) as retrieved using fc7 feature predictions from Q-BOT.

was that if questions are more diverse, the dialog will be more image-informative, and so humans should be able to better guess which image was being talked about. We report top-1 accuracy of true image in human guesses. We also asked humans to rate Q-BOT-A-BOT dialog on consistency, fluency, and level of detail on a 5-point Likert scale.

5 Implementation Details

We used beam search with a beam size of 5 during self-talk between all Q-BOT-A-BOT variants. NDCG scores on the v1.0 val split and the total SL loss (on the same split) were used to select the best SL A-BOT and Q-BOT checkpoints respectively. We used a dropout rate of 0.5 for all SL-pretraining experiments and no dropout for RL-finetuning. We used Adam (Kingma and Ba, 2015) with a learning rate of 10^{-3} decayed by $\sim$0.25 every epoch, upto a minimum of 5×10^{-5}.

The objective for training Diverse-Q-BOT was a sum of the smooth-L1 penalty (introduced in Section 3), cross entropy loss, and L2 loss between the regression head output and the fc7 (Simonyan and Zisserman, 2015) embedding of the image. We observed that coefficients in the range of $1e^{-3}$ to $1e^{-5}$ worked best for the smooth-L1 penalty. We also observed that training for a large number of epochs ($\sim$80) with the above mentioned range of coefficient values led to the best results.

6 Results

- **Q-BOT's diversity** (Table 1): The question-repetition penalty consistently increases diversity (in both SL and RL) over the baseline! `RL: Diverse-Q-bot` asks ~1.5 more unique questions on average than Das et al. (2017b) (6.70 → 8.19) for every 10-round dialog, ~6.3x more novel questions (71 → 449), and a higher fraction and entropy of unique generated n-grams, while still staying comparably relevant (NLL).
- **A-BOT on VisDial** (Table. 2): `RL: A-bot` outperforms `SL: A-bot`, but does not statistically improve over the baseline on answering human questions from VisDial[5] (on v1.0 val & test-std).
- **Image-guessing task** (Fig. 2e): `Diverse-Q-bot + A-bot` (SL and RL) significantly outperform the baseline on percentile rank of ground-truth image as retrieved using Q-BOT's fc7 prediction. Thus, the question-repetition penalty leads to a more informative communication protocol.
- **Human studies** (Fig. 2): Humans judged `RL: Diverse-Q-bot + A-bot` dialog significantly more consistent (fewer conflicting exchanges), fluent (fewer grammatical errors), and detailed (more image-informative) over the baseline and supervised learning. This is an important result. Performance on GuessWhich, together

[5]This is consistent with trends in Das et al. (2017b).

with these dialog quality judgements from humans show that agents trained with our approach develop a more effective communication protocol for the downstream image-guessing task, while still not deviating off English, which is a common pitfall when training dialog agents with RL (Kottur et al., 2017; Lewis et al., 2017).

Note that since our penalty (Eqn. 1) is structured to avoid repetition across successive rounds, one possible failure mode is that Q-BOT learns to ask the same question every alternate dialog round (at t and $t+2$). Empirically, we find that this happens $\sim$15% of times (2490 times out of $\sim$16.5k question pairs) on v1.0 val for RL: Diverse Q-BOT + A-BOT compared to $\sim$22% for SL: Q-BOT + A-BOT. This observation, combined with the fact that Diverse-Q-BOT asks $\sim$1.6 more unique questions relative to SL: Q-BOT across 10 rounds suggests that simply incentivizing diversity in successive rounds works well empirically. We hypothesize that this is because repeating questions every other round or other such strategies to game our repetition penalty is fairly specific behavior that is likely hard for models to learn given the large space of questions Q-BOT could potentially ask.

7 Related Work

Our work is related to prior work in visual dialog (Das et al., 2017a; de Vries et al., 2017) and modeling diversity in text-only dialog (Zhang et al., 2018; Li et al., 2016a,b).

Closest to our setting is work on using conditional variational autoencoders for self-talk in visual dialog (Massiceti et al., 2018), where diversity is not explicitly modeled but is measured via metrics specific to the proposed architecture.

Adding constraints to generate a diverse set of natural language dialog responses have previously been explored in Zhang et al. (2018) via adversarial information maximization, in Gao et al. (2019) by jointly modeling diversity and relevance in a shared latent space, and in Li et al. (2016a) using a maximum mutual information criterion. In contrast, we are interested in diversity at the level of the entire dialog (instead of a single round) – reducing repetitions in QA pairs across multiple rounds. Our repetition penalty is partly inspired by the 'Information Flow' constraint in Li et al. (2016b). As detailed in Section 2, we experimented with similar forms of the penalty and eventually settled on smooth-L1. To the best of our knowledge, we are the first to explicitly model diversity as a constraint in visual dialog.

8 Conclusions & Future Work

We devised an auxiliary objective for training generative Visual Dialog agents with RL, that incentivizes Q-BOT to ask diverse questions. This reduces repetitions in Q-BOT-A-BOT dialog during RL self-talk, and in turn enables A-BOT to be exposed to a larger state space. Through extensive evaluations, we demonstrate that our Q-BOT-A-BOT pair has significantly more diverse dialog while still being image-relevant, better downstream task performance, and higher consistency, fluency, and level of detail than baselines. Our code will be publicly released, and we hope this will serve as a robust base for furthering progress on training visual dialog agents with RL for other multi-agent grounded language games, adapting to learn to talk about novel visual domains, *etc*.

9 Acknowledgements

We thank Nirbhay Modhe and Viraj Prabhu for the PyTorch implementation (Modhe et al., 2018) of Das et al. (2017b) that we built on, and Jiasen Lu for helpful discussions. The Georgia Tech effort is supported in part by NSF, AFRL, DARPA, ONR YIPs, ARO PECASE. AD is supported in part by fellowships from Facebook, Adobe, and Snap Inc. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the US Government, or any sponsor.

References

Samy Bengio, Oriol Vinyals, Navdeep Jaitly, and Noam Shazeer. 2015. Scheduled sampling for sequence prediction with recurrent neural networks. In *NIPS*. 1

Abhishek Das, Satwik Kottur, Khushi Gupta, Avi Singh, Deshraj Yadav, José M.F. Moura, Devi Parikh, and Dhruv Batra. 2017a. Visual Dialog. In *CVPR*. 1, 2, 3, 4, 5

Abhishek Das, Satwik Kottur, José M.F. Moura, Stefan Lee, and Dhruv Batra. 2017b. Learning Cooperative Visual Dialog Agents with Deep Reinforcement Learning. In *ICCV*. 1, 2, 3, 4, 5

Aditya Deshpande, Jyoti Aneja, Liwei Wang, Alexander G Schwing, and David Forsyth. 2019. Fast, di-

verse and accurate image captioning guided by part-of-speech. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 10695–10704. 3

Xiang Gao, Sungjin Lee, Yizhe Zhang, Chris Brockett, Michel Galley, Jianfeng Gao, and Bill Dolan. 2019. Jointly optimizing diversity and relevance in neural response generation. *arXiv preprint arXiv:1902.11205*. 5

Diederik Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *ICLR*. 4

Satwik Kottur, José MF Moura, Stefan Lee, and Dhruv Batra. 2017. Natural language does not emerge 'naturally' in multi-agent dialog. In *EMNLP*. 5

Mike Lewis, Denis Yarats, Yann N Dauphin, Devi Parikh, and Dhruv Batra. 2017. Deal or no deal? end-to-end learning for negotiation dialogues. In *EMNLP*. 5

Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2016a. A diversity-promoting objective function for neural conversation models. 3, 5

Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, and Dan Jurafsky. 2016b. Deep Reinforcement Learning for Dialogue Generation. In *EMNLP*. 3, 5

Daniela Massiceti, N Siddharth, Puneet K Dokania, and Philip HS Torr. 2018. Flipdial: A generative model for two-way visual dialogue. In *CVPR*. 5

Nirbhay Modhe, Viraj Prabhu, Michael Cogswell, Satwik Kottur, Abhishek Das, Stefan Lee, Devi Parikh, and Dhruv Batra. 2018. VisDial-RL-PyTorch. 5

Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. 2016. Sequence level training with recurrent neural networks. In *ICLR*. 1

Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. 2011. A reduction of imitation learning and structured prediction to no-regret online learning. In *AISTATS*. 1

Iulian Vlad Serban, Alberto García-Durán, Çaglar Gülçehre, Sungjin Ahn, Sarath Chandar, Aaron C. Courville, and Yoshua Bengio. 2016. Generating Factoid Questions With Recurrent Neural Networks: The 30M Factoid Question-Answer Corpus. In *ACL*. 3

Karen Simonyan and Andrew Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. In *ICLR*. 3, 4

Harm de Vries, Florian Strub, Sarath Chandar, Olivier Pietquin, Hugo Larochelle, and Aaron Courville. 2017. GuessWhat?! visual object discovery through multi-modal dialogue. In *CVPR*. 1, 5

Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256. 1, 2

Yizhe Zhang, Michel Galley, Jianfeng Gao, Zhe Gan, Xiujun Li, Chris Brockett, and Bill Dolan. 2018. Generating informative and diverse conversational responses via adversarial information maximization. In *NIPS*. 3, 5