

AQM+: A new model for visual dialog question generation

March 8 2019, by Ingrid Fadelli





Credit: Lee et al.

Researchers at Clova AI Research, NAVER and LINE, have recently proposed a new framework called AQM+ that allows dialog systems to generate context-relevant questions and answers. Their model, outlined in [a paper pre-published on arXiv](#), will be presented at the 7th International Conference on Learning Representation ([ICLR 2019](#)), in New Orleans.

"Intra-machine and human-machine collaboration has been considered as a significant and meaningful research topic, in particular, from the perspective of ethics and public interest in AI," Sang-Woo Lee, one of the researchers who carried out the study, told TechXplore. "Focusing on

task-oriented dialog (TOD), researchers have gained considerable insight from GO games between humans and AlphaGo. More specifically, these researchers think that goal-oriented dialog models can be improved by training models on giant-scale machine-machine interactions, which allows AlphaGo to beat human experts. However, I do not totally agree with this idea, because dialog is a task based on collaboration between two players and is fundamentally different from Go, which is a competition game."

According to Lee, task-oriented dialogue (TOD) resembles the game 20 Questions, as the dialog system's goal should be that of narrowing down a user's intents and requests. In [a previous study](#), Lee and his colleagues introduced a framework for task-oriented dialog systems called "answerer in questioner's mind" (AQM) that is based on this idea. AQM allows dialog systems to ask questions that maximize their information gain, reducing the uncertainty of the user's intent.

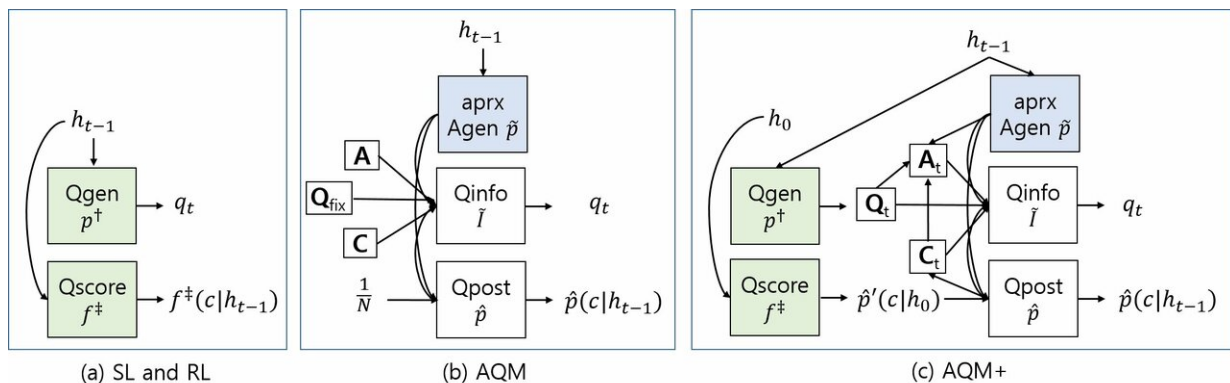
Image + Caption	SL-Q	AQM+ w/ indA
 <p>a bird is standing on a rock in the water</p>	<p>are there any people? no</p> <p>what color is water? blue</p> <p>what color is water? blue</p> <p>is it sunny? yes</p> <p>are there any people? yes</p> <p>are there any other animals in picture? no</p> <p>is it 's sky? yes</p> <p>are there clouds? ca n't ell</p> <p>is bird 's beak? yes</p>	<p>are there any people? no</p> <p>what kind of bird is it 's body? falcon maybe , it 's small</p> <p>what color of parrot? green</p> <p>are there leaves? yes</p> <p>what color is bird? black and white</p> <p>what color of water? blue</p> <p>does it is water? yes</p> <p>are there trees in ocean? yes</p> <p>are there grass? yes</p> <p>can you see sky? no</p>
	AQM+ w/ indA (gen1Q, Q = 100)	
	<p>what color is there is there? it is black and white</p> <p>what color is sky? ca n't see it</p> <p>what kind of lake? i ca n't tell</p> <p>what kind of lake? i ca n't tell</p> <p>what kind of lake? i ca n't tell</p> <p>what kind of lake? i ca n't tell</p> <p>what kind of lake? i ca n't tell</p> <p>is there grass? i ca n't see any</p> <p>what color is stream? it 's light brown</p> <p>what kind of water? it looks like stream</p>	<p>what is man wearing? i can only see his shirt which is blue</p> <p>any words on bus? yes , it says UNK</p> <p>is anything in vase? yes , it is green</p> <p>can you tell what time of day it is? no</p> <p>are cows on farm? i ca n't tell , but i think so</p> <p>any words on bus? yes , it says UNK</p> <p>what is man wearing? i can only see his shirt which is blue</p> <p>any words on bus? yes , it says UNK</p> <p>is anything in vase? no</p> <p>does man look happy? yes</p>
RL-QA		
	<p>how many birds are there? i see 6 seagulls and it would say that it is that it is that caption they are real picture</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it caption caption is</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it that caption</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it caption</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it caption</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it that it is</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it is it</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it is it</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it is it</p>	
Image + Caption	SL-Q	AQM+ w/ indA
 <p>a pizza sitting on top of a white plate</p>	<p>is it is box on plate? yes</p> <p>what color is plate? white</p> <p>is it large pizza? yes</p> <p>is there appear to cut? no</p> <p>is this look like drinks? no</p> <p>is this on plate? yes</p> <p>is there is there? no</p> <p>is this is there? no</p> <p>is there is restaurant? no</p> <p>is there is pizza? yes</p>	<p>are there see any people? no</p> <p>what kind of restaurant? ca n't tell</p> <p>any other than pizza on? no</p> <p>any silverware? no</p> <p>what color of pizza? white and green</p> <p>what is on pizza? cheese</p> <p>are there is table? yes</p> <p>is table have table? no</p> <p>are there anything else? no</p> <p>what else can you? there is pizza in background</p>
	AQM+ w/ indA (gen1Q, Q = 100)	
	<p>are there see any people are there? no</p> <p>is plate have restaurant? i ca n't tell</p> <p>is plate have restaurant? i ca n't tell</p> <p>is plate have restaurant? i ca n't tell</p> <p>is pizza on table? i ca n't tell</p> <p>is there see pizza in image? no just plate</p> <p>is pizza on table? i ca n't tell</p> <p>is pizza on table? i ca n't tell</p> <p>is pizza on table? i think so</p> <p>is pizza on table? i think so</p>	<p>is anyone on sofa? no</p> <p>is it guy or girl? it is man</p> <p>is floor carpeted? no</p> <p>are mushrooms sliced thinly or wide piece? they are cut into slices</p> <p>can you tell if cake is chocolate? it is chocolate frosted donut</p> <p>how does weather look? i ca n't really tell</p> <p>is it guy or girl? it is boy</p> <p>can you tell if cake is chocolate? it is chocolate frosted</p> <p>how does weather look? it looks like it is sunny</p> <p>is it boy or girl? it is boy</p>
RL-QA		
	<p>do you see any people? no , just pizza is wrong no people around it is table is caption it is no i would call</p> <p>can you tell what room they are? i ca n't , no it is it is that it is table it is caption is it caption</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it caption is that it</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is average picture , no ,</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is that it is</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is no it is it caption caption</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is that it is that</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it is it</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it is it</p> <p>can you tell what kind of water is? no , it is it is it is it is it is it is it is it is it is it is it</p>	

Credit: Lee et al.

Unlike other approaches, AQM explicitly calculates the posterior distribution and finds solutions analytically. Despite its many advantages, AQM was found to generalize poorly in real-world tasks, where the number of objects, questions and answers are typically unrestricted.

In their recent study, the researchers addressed this limitation and proposed a new approach, AQM+. Unlike their previous approach, AQM+ can be applied to large-scale problems, generating questions and answers that are more coherent with the changing context of a given dialog.

"Similar to human dialog, our AQM+ models what the opponent says and reasons the most effective strategy to grasp the opponent's mind and intent, using an information theory metric (information gain)," Lee explained. "This approach is different from previous neural network-based methods for TOD, which mainly employed sequence-to-sequence (Seq2Seq) for directly generating questions by responding to the previous utterance."

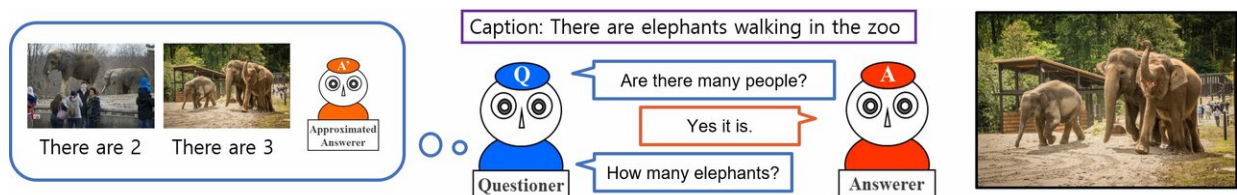


Credit: Lee et al.

Lee and his colleagues evaluated AQM+ on a challenging task-oriented visual dialog problem called GuessWhich. Their model achieved remarkable results, outperforming state-of-the-art approaches by a considerable margin.

"The approach based on our 20 Questions game in AQM+ for questioning users can tackle complex dialog situations where there exist many and various answers and cases related to general-formed questions, as well as yes or no questions," Lee said. "This means that our AQM+ can be applied to different TOD situations in the real world."

In their tests, Lee and his colleagues Jung-Woo Ha, Tong Gao, Sohee Yang and Jaejun Yoo found that AQM+ reduced errors by 60 percent as a dialog proceeds, while existing algorithms achieved an error reduction of less than 6 percent. According to the researchers, AQM+ could be used to generate both open and closed questions.



Credit: Lee et al.

"Effectively training models from dialog data in an end-to-end manner remains highly challenging, particularly for the development of TOD systems," Jung-Woo Ha, another researcher involved in the study, told TechXplore. "Although AQM+ mainly focuses on questioning to get useful information from the answerer, it can be naturally extended by

including answering the questions based on the same approach."

Lee, Ha and their colleagues are now considering a number of directions for future research. Firstly, they would like to develop their approach further in order to achieve a general learning framework for dialog. Their ultimate objective is to design a system that can achieve human-like accuracy in communicating with humans.

"Ultimately, we aim to develop a general AI framework that enables human-like machine-machine and machine-human dialogs," Ha said. "As industrial research scientists, we will apply our technologies to diverse services such as messenger and AI assistant platform, thus offering greater value for global users."

More information: Large-scale answerer in questioner's mind for visual dialog question generation. arXiv:1902.08355 [cs.CL].
arxiv.org/abs/1902.08355

Answerer in questioner's mind: information theoretic approach to goal-oriented visual dialog. arXiv:1802.03881 [cs.CV].
arxiv.org/abs/1802.03881

© 2019 Science X Network

Provided by Science X Network

Citation: AQM+: A new model for visual dialog question generation (2019, March 8) retrieved 26 April 2024 from <https://techxplore.com/news/2019-03-aqm-visual-dialog.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.