ASK THE RIGHT QUESTION

Active Question Reformulation with Reinforcement Learning

SNU Datamining Laboratory
2018. 2. 27 Seminar
Sungwon, Lyu
lyusungwon@dm.snu.ac.kr

QUESTION ANSWERING

- Answering a question based on context
- Datasets
  - bAbI dataset, Deepmind QA dataset, SQuAD dataset
- Models
  - Memory Network, End-to-End Memory Network, Gated End-to-End Memory Network, Dynamic Memory Network, Dynamic Memory Network +, Relation Memory Network

Source: Large-scale simple question answering with memory networks (Weston, J., Chopra, S., & Bordes, A., ICLR 2015), End-to-End Memory Network (Sukhbaatar, S., Weston, J., & Fergus, R., NIPS 2015), Gated End-to-End Memory Network (Liu, F., & Perez, J., ACL 2017, Ask me anything: Dynamic memory networks for natural language processing (Kumar, A., et al., ICML 2016), Dynamic memory networks for visual and textual question answering (Xiong, C., Merity, S., & Socher, R., ICML 2016), Relation Memory Network: A Simple Neural architecture for Text-based Reasoning (Jihyung Moon, 2018)

SEARCHQA DATASET

- Dataset of noisy information retrieval environment
- QA from TV show 'Jeopardy!' + Google Search results of each question
- 140,461 Question - Answer sets each coupled with 49.6 snippets on average
- Each snippet is 37.3 tokens / Answer is 1.47 tokens on average
- Example
  - Question: Travel doesn't seem to be an issue for this sorcerer & onetime surgeon; astral projection & teleportation are no prob
  - Snippet1: Astral Projection - Spells - D&D Beyond You and up to eight willing creatures within range project your astral bodies into the Astral Plane (the spell fails and the casting is wasted if you are already on that plane). The material body you leave behind is unconscious and in a state of suspended animation; it doesn't need food or air and doesn't age. Your astral body …
  - Answer: Doctor Strange
- Result

Answer	Unigram	n-gram
Per-question Average	66.97%	42.86%
Per-user Average	64.85%	43.85%
Per-user Std. Dev.	8.16%	10.43%
F1 score (for n-gram)	-	57.62 %

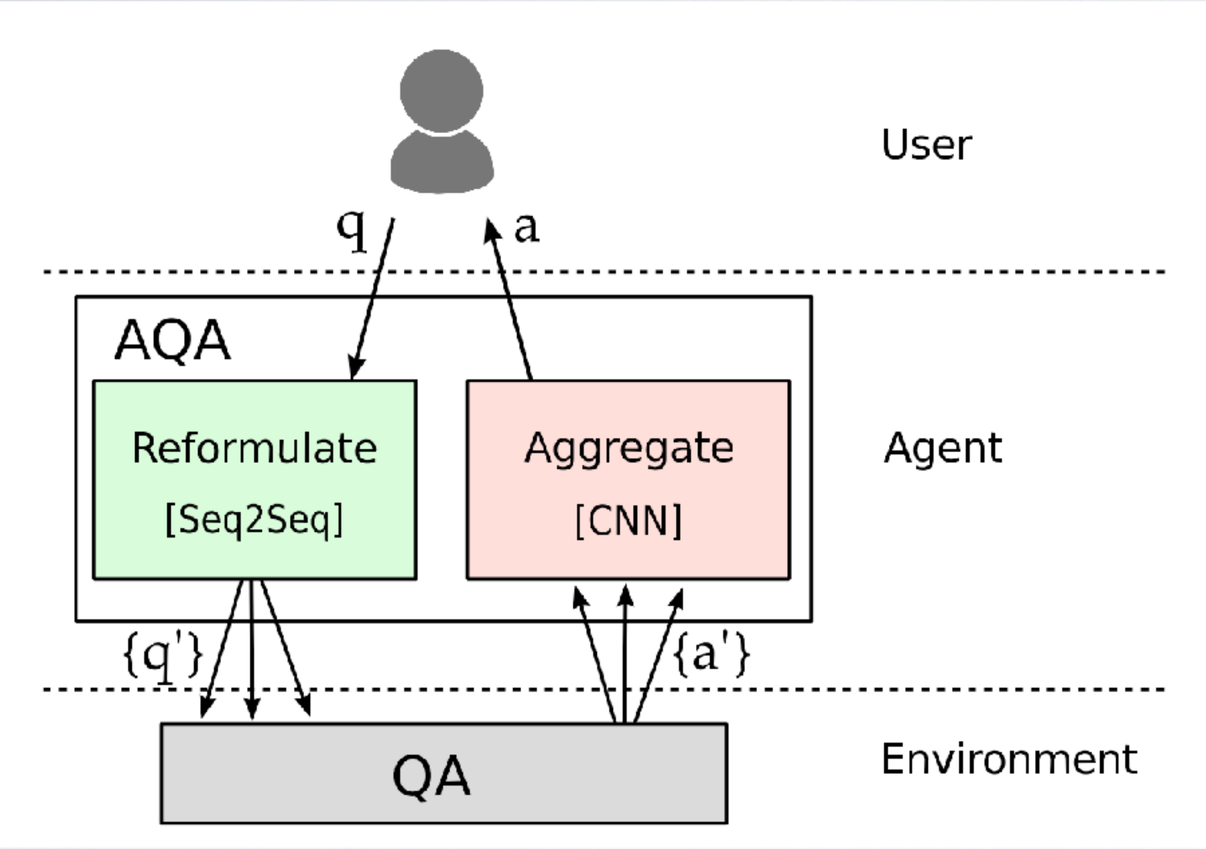
Model	Set	Unigram Acc	Unigram Acc@5	n-gram F1
TF-IDF	Valid	13.0	49.3	–
Max	Test	12.7	49.0	–
ASR	Valid	43.9	67.3	24.2
	Test	41.3	65.1	22.8

Source: SearchQA (Dunn et al, 2017)

MOTIVATION


Google
Bugles
Advanced Search
Language Tools
Where are the **Bugles**?
Are **Bugles** here to stay?
Is the **Bugles** factory near a graveyard?
Who dreams of the **Bugles**?
What happens to **Bugles** after I look at them?
What are the worst operas about **Bugles**?
It is good to eat too many **Bugles**?
Who invented **Bugles** and is he dead?
What are the longest movies about looking at bags of **Bugles**?
How many **Bugles** would it take to sink a cruise ship?
Are **Bugles** the most delicious thing that soldiers have ever eaten?
What is the **Bugles** shape called?
Do I belong with the **Bugles**?
Google Search
I'm Feeling Lucky

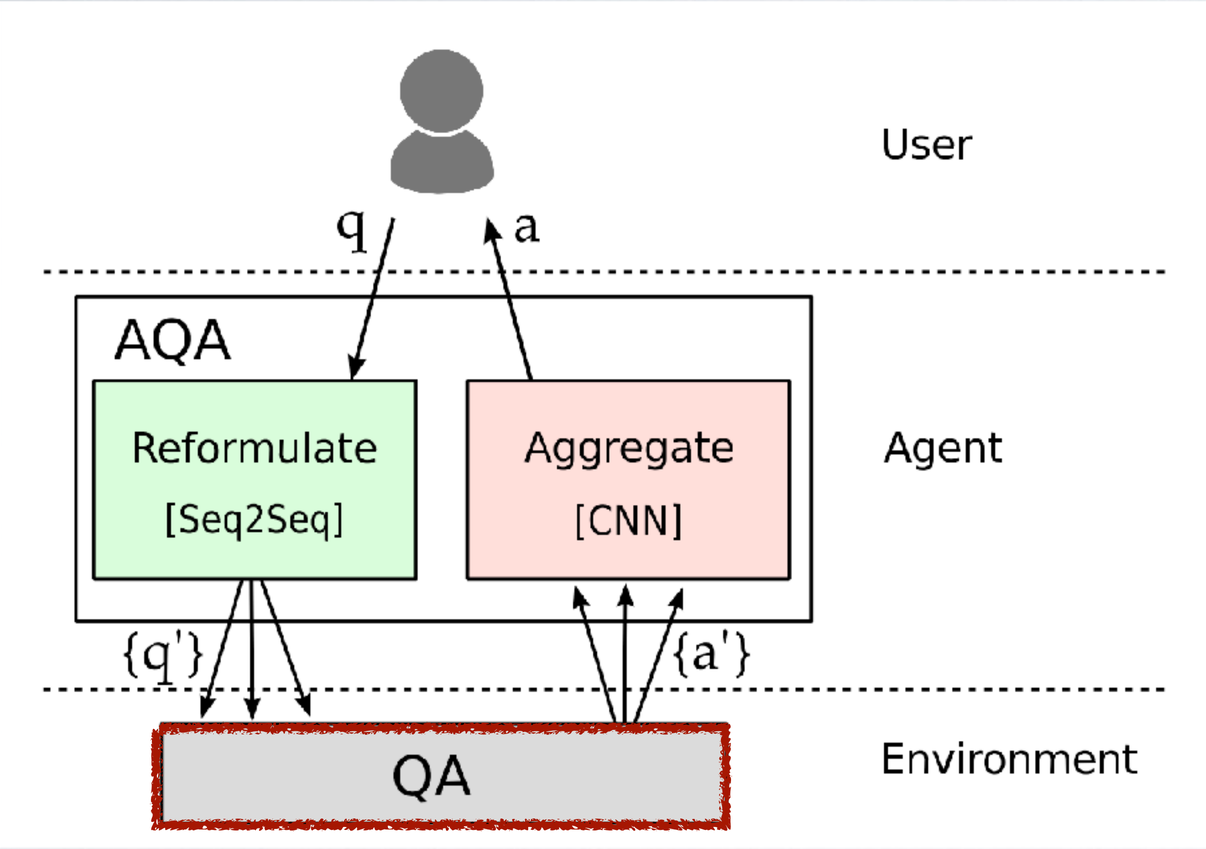
AGENT - ENVIRONMENT FRAMEWORK



- Question-Answering Environment
  - Return an answer given a question and snippets set
- Reformulation Model
  - Sequence to sequence monolingual paraphrasing model
  - Use F1 score of the answer as reward given a reformulated question
- Answer selection model (for test)
  - Choose answer from several question - rewrite - answer sets

Source: Ask the right questions: Active question reformulation with reinforcement learning (C. Buck et al, ICLR 2018)

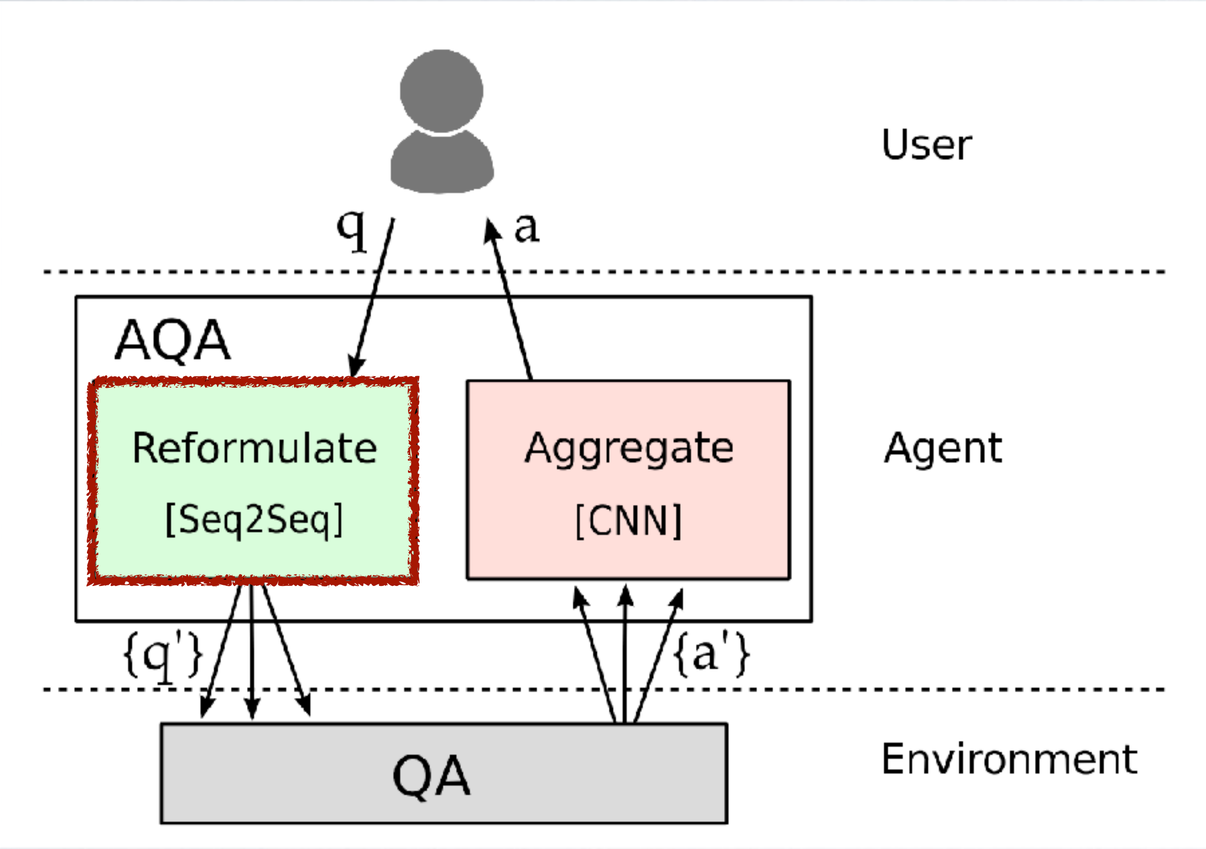
TRAINING



- Question-Answering Environment
  - BiDirectional Attention Flow(BiDAF)
    - extractive QA system
  - Used top 10 snippets
  - Adam, 4500 steps, lr 0.001, batch 60

Source: Bidirectional attention flow for machine comprehension (M Seo et al, ICLR 2017)

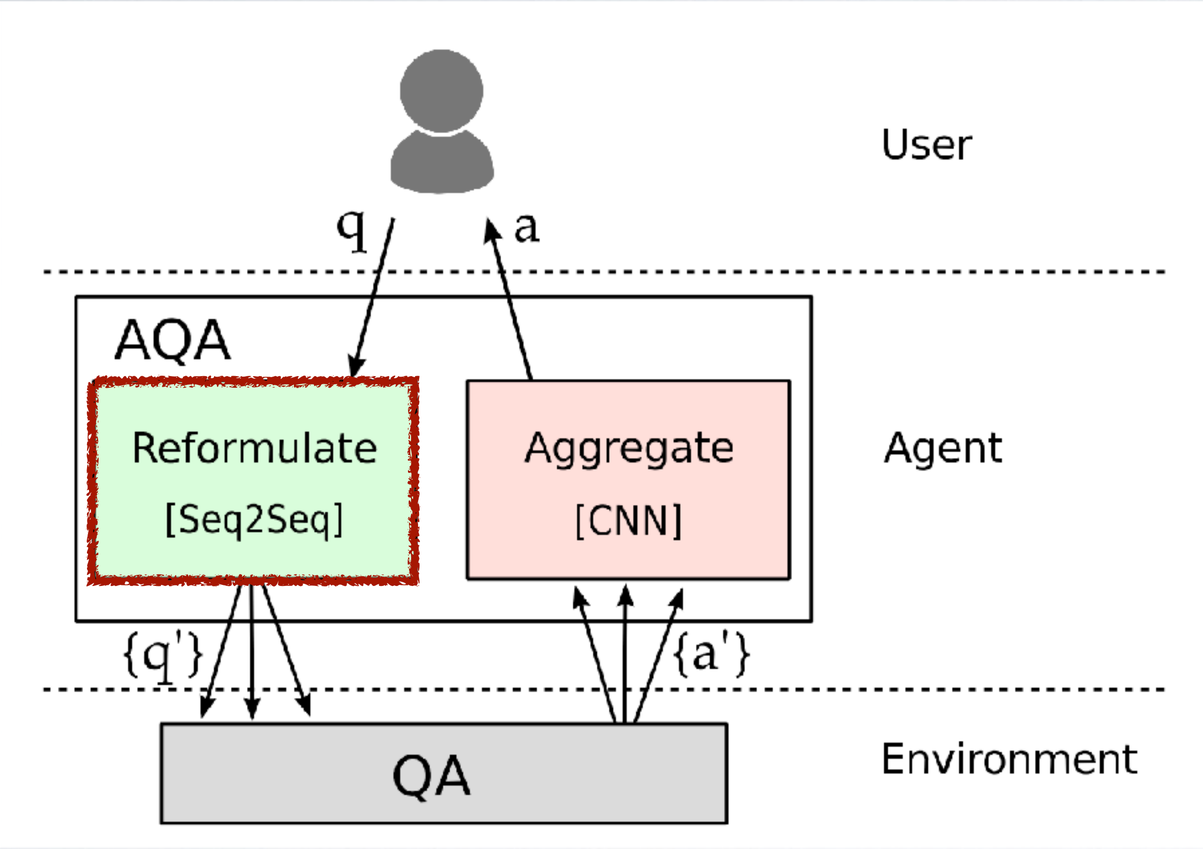
TRAINING



- Reformulation Model (Pretrain)
  - Google Multilingual NMT (zero-shot)
    - bidirectional LSTM as encoder
    - 4layers LSTM as decoder
  - UN parallel corpus data (30 languages)
  - Paralex monolingual data (1.5M pairs)
  - 400M instances, Adam, lr 0.001, batch 128

Source: Google's neural machine translation system: Bridging the gap between human and machine translation (Y Wu et al, ICLR 2017), Google's multilingual neural machine translation system: enabling zero-shot translation (M Johnson et al, 2016)

TRAINING



- Reformulation Model (Reinforcement)
  - Policy Gradient

$$\pi_\theta(q|q_0) = \prod^{T} p(w_t|w_1, \ldots, w_{t-1}, q_0)$$

$$\mathbb{E}_{q\sim\pi_\theta(\cdot|q_0)}[R(f(q))] \approx \frac{1}{N}\sum_{i=1}^{N} R(f(q_i)), \quad q_i \sim \pi_\theta(\,\cdot\,|q_0)$$

$$\nabla\mathbb{E}_{q\sim\pi_\theta(\cdot|q_0)}[R(f(q))] = \mathbb{E}_{q\sim\pi_\theta(\cdot|q_0)}\nabla_\theta \log(\pi_\theta(q|q_0))R(f(q))$$

$$\approx \frac{1}{N}\sum_{i=1}^{N} \nabla_\theta \log(\pi(q_i|q_0))R(f(q_i)), \quad q_i \sim \pi_\theta(\,\cdot\,|q_0)$$

$$B(q_0) = \mathbb{E}_{q\sim\pi_\theta(\cdot|q_0)}[R(f(q))]$$

  - Entropy Regularization

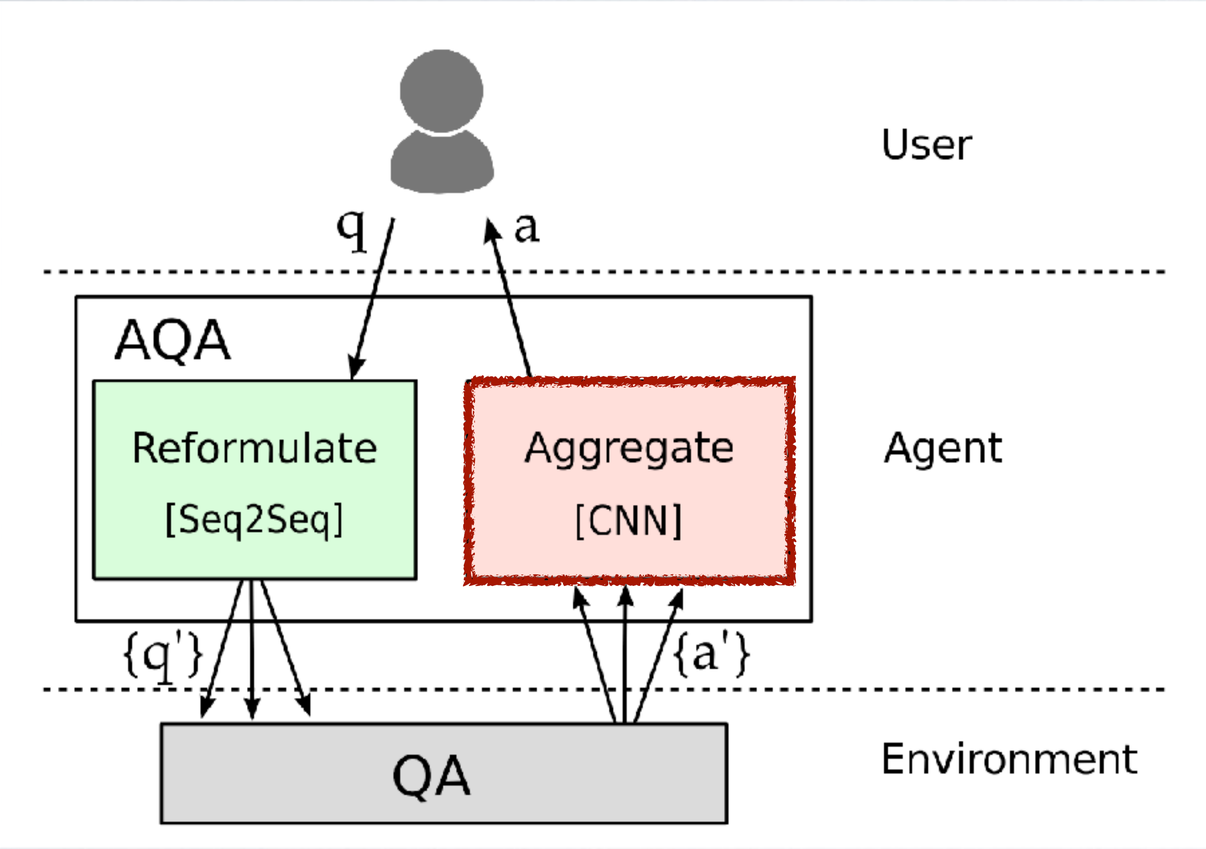
$$H[\pi_\theta(q|q_0)] = -\sum_{t=1}^{T}\sum_{w_t\in V} p_\theta(w_t|w_{<t}, q_0)\log p_\theta(w_t|w_{<t}, q_0)$$

  - Objective Function

$$\mathbb{E}_{q\sim\pi_\theta(\cdot|q_0)}[R(f(q)) - B(q_0)] + \lambda H[\pi(q|q_0)]$$

  - 100k steps, lr 0.001, rr 0.001, batch 64

TRAINING



- Answer Selection Model
  - 1D CNN + max pooling, window size 3
  - Train to predict whether the F1 score of Question - rewrite -answer pair is above or below average (binary classification)
  - Work as ensemble

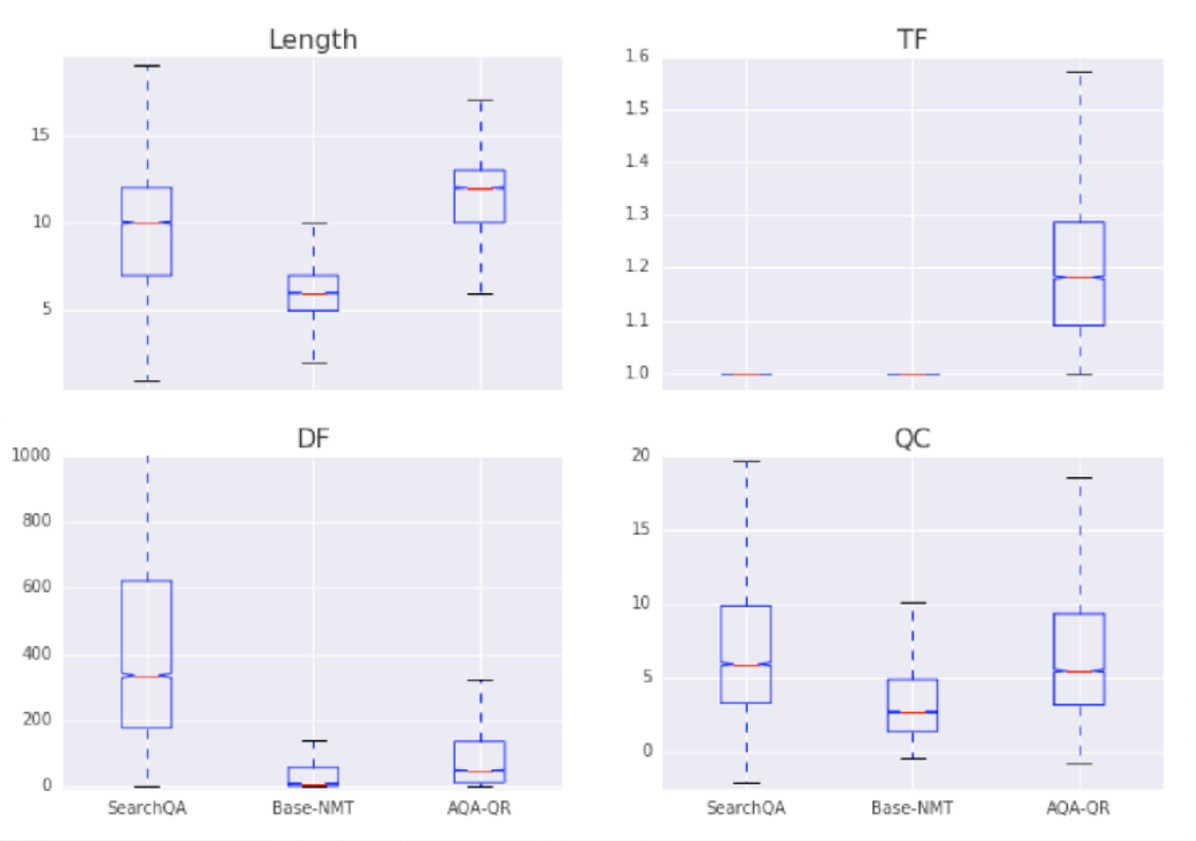
RESULT

Model	Query	Reference / Answer from BiDAF (F1)
Jeopardy!	People of this nation AKA Nippon wrote with a brush, so painting became the preferred form of artistic expression	japan
SearchQA	people nation aka nippon wrote brush , painting became preferred form artistic expression	japan (1.0)
MI	nippon brush preferred	julian (0)
Base-NMT	Aka nippon written form artistic expression?	julian (0)
AQA-QR	What is name did people nation aka nippon wrote brush expression?	japan (1.0)
AQA-Full	people nation aka nippon wrote brush , painting became preferred form artistic expression	japan (1.0)

RESULT

		Baseline		MI-SubQuery		Base-NMT		AQA				
		ASR	*BiDAF*	*TopHyp*	*CNN*	*TopHyp*	*CNN*	*TopHyp*	*Voting*	*MaxConf*	*CNN*	*Human*
Dev	EM	-	31.7	24.1	37.5	26.0	37.5	32.0	33.6	35.5	**40.5**	-
	F1	24.2	37.9	29.9	44.5	32.2	44.8	38.2	40.5	42.0	**47.4**	-
Test	EM	-	28.6	23.2	35.8	24.8	35.7	30.6	33.3	33.8	**38.7**	43.9
	F1	22.8	34.6	29.0	42.8	31.0	42.9	36.8	39.3	40.2	**45.6**	-

ANALYSIS



- Search QA
  - Gandhi was deeply influenced by this count who wrote "War and Peace" is simplified to gandhi deeply influenced count wrote war peace.
- Base NMT
  - Who influenced count wrote war?
- AQA-QR(Question Reformulator)
  - What is name gandhi gandhi influence wrote peace peace?

IMPLICATION

- Reinforcement framework is useful in in-differentiable environment.
- Models may have their own suitable languages
- AQA may be useful in other environments of information retrieval
- Reformulate anything?

Source: Task-Oriented Query Reformulation with Reinforcement Learning (R Nogueira et al, 2017)