## Supplementary File: TVQA: Localized, Compositional Video Question Answering

Jie Lei Licheng Yu Mohit Bansal Tamara L. Berg Department of Computer Science University of North Carolina at Chapel Hill

{jielei, licheng, mbansal, tlberg}@cs.unc.edu

## A Supplemental Material

## **Questions and Answers:**

Fig. 1 shows the distribution of question lengths. Most of our questions are longer than 10 words. Fig. 2 shows a pie distribution of our first-part question words, as well as the associated second-part question words. We see that 'what' accounts for a large portion for the first part, i.e., what (Abstract, Object, Action), and 'when' is more often used for moment localization in the second part. Fig. 3 shows the distribution of answer lengths. Note there is no peak answer length in TVQA, thus understanding the variable-length answers is also crucial for TVQA question-answering.



Figure 1: Distribution of question lengths.



Figure 2: Distribution of questions by their first words and link words.

**More QA examples:** We show more examples using our S+V+Q joint model trained on full-length video and subtitle in Fig. 4, where the top three



Figure 3: Distribution of answer lengths.

rows show the correct answers made by our model and the bottom two rows show the wrong ones.

Character	Top unique nouns
Derek	Anna, MRI, today, tumor, black, desk,
	car, area, trailer, procedure, ice
Cristina	Colin, dress, leg, Doug, Deborah, wedding,
	choker, restroom, lab, pad
Izzie	Cheyenne, check, Jilly, sweater, Heather,
	bathroom, lunch, sheet, beast, condition
Burke	surgeon, valve, pole, cap, fishing, point,
	situation, look, flower, pressure, breakfast
Meredith	Susan, thatcher, mom, counter, Dylan,
	boyfriend, band, sister, step, Molly
George	mackie, Sophie, conference, west, mercy,
	box, stairwell, pocket, bottle, morning
Bailey	Viper, son, clinic, clipboard, gesture, binder,
	Kelley, razor, scalpel, sex, piece, twin, Judy
Addison	Naomi, ring, Sam, news, rose, Lisa, station,
	kid, ultrasound, pink, monitor, concern

Table 1: Top unique nouns for characters in Grey.

Character	Top unique nouns
Ted	class, girlfriend, tattoo, movie, architect,
	Jen, word, Hammond, airport, student
Lily	gift, purse, dress, kitchen, water, card, tree,
	plan, Judy, credit, mouth, paint, bedroom
Barney	James, Jerry, Nora, suit, brother, Bob,
	Sam, strip, cigar, way, fist, street, tie
Marshall	sandwich, brad, burger, Jeff, money,
	song, work, Marvin, laptop, school,
Zoey	landmark, thanksgiving, cousin, tape,
	husband, spot, bit, turkey, text, recording

Table 2: Top unique nouns for characters in HIMYM.

**Top unique nouns for characters:** We show the top unique unique nouns for characters for Grey and HIMYM in Tab. 1 2.



(Alvie:)Yo, I don't need an escort. Get off me. 00:26.382 --> 00:28.612 (Alvie:)Who are you? You believe these guys?

What does Alvie say when he is being brought into the room with House ?

a0 Go Away a1 Get Off Me √ a2 Leave me alone a3 Goodbye a4 Hello

(a)



01:10.131-->01:12.292 (Meredith:)I was swimming. 01:14.358 (Meredith:)I was fighting. (Meredith:)And then I thought, just for a second.

What was behind Meredith when she said she was swimming?

a0 A painting a1 A couch a2 A metal shelf√ a3 A file cabinet a4 A car

(c)



00:11.332 --> 00:12.958 (Leonard:)Aw. (Penny:)So handsome. > 00.16.83 (Penny:)Like James Bond. \_\_\_\_\_ What was Leonard wearing when he walked into the room with Amy ,

Penny and Sheldon ? a0 A button down shirt with a vest a1 A pinstriped suit a2 A track suit tuxedo √ a4 A clown costume

(b)

00:00.566 --> 00:03.700 (Sheldon:)I'm sensing awkwardness, am I right? 00:03.70 > 00:05.669 (Howard:)Yes. 00.07737 0.05.67 (Sheldon:)Swish.

Where is Amy when Sheldon asks her if he is sensing awkwardness?

a0 Amy is sitting on Penny 's lap a1 Amy is sitting on Howard 's lap a2 Amy is sitting on the sofa next to Sheldon √ a3 Amy is standing behind Penny's chair a4 Amy is leaning against the front door

(d)



-> 00:47.875 00:43 (Callie:)When you put your freshly severed finger into a grungy, bacteria-filled glove, 00:47.948-->00:51.042 (Callie:)- you got a severe infection.

What was wrong with Heath 's hand before he put it into the glove ?

a0 He couldn't feel his fingers al His finger had been freshly severed v a2 His hands were cold



00:12.27 > 00:13.640 (Derek:)Cristina write her vows? (Burke:)I guess I'll find out tomorrow. (Burke:)- How's Meredith doing?

Who does Burke ask Derek about when they are eating peanuts?

a0 Burke asks about James al Burke asks about Meredith a2 Burke asks about Addison

a3 Burke asks about Addison a4 Burke asks about Bailey

(f)



0.06.426 >00.0777 (Captain:)of my betrayer is ... (Captain:)when I tell you who it is, you're gonna .... > 00.1 (Captain:)It's my doorman!

What is captain holding in his hand when he tells Ted that he believes his doorman is his betrayer a0 A mace and chain  $\sqrt{}$ al A snake a2 An Ax a3 A sword a4 A gun



(Sheldon:)Excuse me, Penny, but Doodle Jump is a game. (Sheldon:)Angry Birds is a game. (Sheldon:)World of Warcraft is a ...

How did Sheldon wear his watch when he was talking to Penny?

a0 He was wearing it on his left wrist al He was wearing it on his right wrist  $\sqrt{}$ a2 He didn't have a watch a3 He was wearing it on his left ankle a4 He was wearing it on his left ankle

(h)

01:02.935-->01:07.372 (Derek:)Follow my finger. Up, down. Side to side... 01:07.439 --> 01:10.23 (Derek:)Good. Let me know if you can feel this. OK?

How many men were in the beds in the room when Derek and Preston were examining them ? a0 4 a1 2  $\checkmark$  a2 5

(j)

a3 3 a4 6

Figure 4: More examples from TVQA dataset, along with the predictions from our best model. Ground truth answers are highlighted in green, and model predictions are indicated by  $\checkmark$ . Best viewed in color.



a3 All of Heath 's fingers were cut off a4 He had been shot in the hand



(g)

00:00.330 --> 00:02.382 I guess this chair's mine now

(Chandler:)Joey, you broke my chair! 00:07.532 --> 00:08.468

00:07.532 --> 00.00.400 (Rachel:)Your chair?

What does Joey do when he said he guesses this chair was his ?

00:05.723 --> 00:07.276

a0 He sits on the chair a1 He stands up  $\sqrt{}$ a2 He walks over to Chandler a3 He picks up the phone a4 He picks up a sandwich

(i)

