From: "Zündorf, Monica-Laura (IAR)" <monica-laura.haurilet@kit.edu>
To: Tomasz Kornuta <tkornut@us.ibm.com>

Date: Tuesday, October 08, 2019 09:25AM
Subject: [EXTERNAL] Re: It’s not about the Journey... its the results ;)

For Follow up: Normal Priority.

History: This message has been replied to and forwarded.

Hi Tomasz,

I am well. I hope you too :)

The conference went well. I also really liked Sydney.

I was able to generate the task-level performance of our model.

So sorry that it took so long... I wanted to run it yesterday but I had to work on a paper (the deadline is today...). :(

Here are the numbers:

cog_AndSimpleExistShapeGo.json.gz: 100.00
cog_GoColor.json.gz: 100.00
cog_SimpleCompareColorGo.json.gz: 93.99
cog_Exist.json.gz: 98.64
cog_ExistShapeSpaceGo.json.gz: 97.92
cog_SimpleCompareColor.json.gz: 94.48
cog_GoShape.json.gz: 100.00
cog_CompareShape.json.gz: 95.11
cog_ExistSpaceGo.json.gz: 98.55
cog_AndSimpleExistColorGo.json.gz: 99.97
cog_SimpleCompareShape.json.gz: 94.08
cog_GoColorOf.json.gz: 100.00
cog_GetShape.json.gz: 100.00
cog_ExistShapeOf.json.gz: 99.79
cog_ExistColorSpace.json.gz: 97.78
cog_GetColor.json.gz: 100.00
cog_ExistSpace.json.gz: 98.63
cog_ExistColor.json.gz: 100.00

cog_AndSimpleExistGo.json.gz: 100.00
cog_ExistLastShapeSameColor.json.gz: 100.00
cog_AndSimpleCompareShape.json.gz: 93.40
cog_ExistShapeGo.json.gz: 100.00
cog_GoShapeOf.json.gz: 100.00
cog_GetShapeSpace.json.gz: 98.29
cog_CompareShapeGo.json.gz: 95.52
cog_SimpleCompareShapeGo.json.gz: 93.97
cog_SimpleExistGo.json.gz: 100.00
cog_ExistLastObjectSameObject.json.gz: 100.00
cog_ExistShape.json.gz: 100.00
cog_AndCompareShape.json.gz: 76.98
cog_CompareColorGo.json.gz: 95.07
cog_Go.json.gz: 100.00
cog_ExistShapeSpace.json.gz: 97.95
cog_ExistColorSpaceGo.json.gz: 97.75
cog_ExistLastColorSameShape.json.gz: 100.00
cog_SimpleExistShapeGo.json.gz: 100.00
cog_ExistColorGo.json.gz: 100.00
cog_AndSimpleCompareColor.json.gz: 92.87
cog_ExistGo.json.gz: 100.00
cog_SimpleExistColorGo.json.gz: 100.00
cog_CompareColor.json.gz: 95.65
cog_AndCompareColor.json.gz: 77.86
cog_GetColorSpace.json.gz: 98.27
cog_ExistColorOf.json.gz: 99.72

Best,

Monica

From: Tomasz Kornuta <tkornut@us.ibm.com>
Sent: Thursday, October 3, 2019 12:10 AM
To: Zündorf, Monica-Laura (IAR)
Subject: RE: It’s not about the Journey... its the results ;)

Hi Monica,

Hope you are doing well. Just poking you once again about the per-task results. ;)

All the best,
Tomasz

> ----- Original message -----
> From: "Zündorf, Monica-Laura (IAR)"<monica-laura.haurilet@kit.edu>
> To: Tomasz Kornuta <tkornut@us.ibm.com>

Cc:
Subject: [EXTERNAL] Re: It’s not about the Journey... its the results ;)
Date: Fri, Sep 20, 2019 2:04 AM

Hello Tomasz,

yes, the results are on the Canonical setting. We haven't evaluated on the hard variant and, unfortunately, I will not have time to try it out.

Currently I am at a conference (+holiday) and won't be able to give you more results. In 2 weeks I will be back at the lab and then I will try to calculate the per-task accuracies. :)

I hope it is early enough for you?

Have a great day,

Monica

--

Monica Haurilet
PhD Student
Computer Vision for Human-Computer Interaction Lab
Karlsruhe Institute of Technology (KIT), Germany
https://cvhci.anthropomatik.kit.edu/~mhaurile/

From: Tomasz Kornuta <tkornut@us.ibm.com>
Sent: Thursday, September 19, 2019 6:16 PM
To: Zündorf, Monica-Laura (IAR)
Subject: It’s not about the Journey... its the results ;)

Hey,

My name is Tomasz and I am a research scientist working in IBM Research. We are focusing on Visual Reasoning and in particular we are working on COG. We have found your paper about SoftPaths interesting and would like to compare your results with the results achieved by our model. Sadly, we couldn't find detailed performance (accuracies) "per task" anywhere (paper, supplemental, poster).

As we focus currently only on answering task (we skip pointing) we could simply "average" the results from Table 2 (exclude pointing and average existence, conditional questions and questions about object attributes), but want to look at each individual task. Besides, I assume those are results on Canonical variant of COG,

> right? Do you have results for the Hard variant?
>
> We would be grateful if you could sent us the table or sth with per-task accuracies of your best model. If you have them for both Canonical and Hard - that would be even better! ;)
>
> Thanks,
> Tomasz Kornuta
> IBM Research AI
> Almaden Research Center, CA, US