

R-C²: Cycle-Consistent Reinforcement Learning Improves Multimodal Reasoning

Supplementary Material

A. Training Hyperparameters

The key global hyperparameters used throughout all experiments in this paper are listed in Table A1. Unless explicitly stated otherwise, all models were trained under this shared configuration.

Parameter	Value
Global Batch Size	256
Learning Rate	1×10^{-6}
Weight Decay	1×10^{-2}
KL Penalty Coefficient (λ_{KL})	1×10^{-2}
Max Steps	100
Number of Rollouts	4
Temperature	1.0
Top-p	0.95

Table A1. Key global hyperparameters used throughout our experiments.

B. Reproposed VWA Task

To enable a more straightforward evaluation of multimodal web understanding, we redesign the VWA dataset into a multiple-choice question answering (MC-QA) task based on shopping-website pages. For each webpage, we automatically generate question-answer pairs using GPT-5, followed by human verification to ensure correctness and quality. Each page contains **10** questions, and each question provides **6** candidate choices, with **1** correct answer. Visualized examples are provided in Figures 1 and 2.

C. Prompt Templates


We list the prompt templates used during training and evaluation in Table A3.

My Account My Wish List Sign In Welcome to One Stop Market Create an Account


One Stop Market Search entire store here... Advanced Search

Beauty & Personal Care · Sports & Outdoors · Clothing, Shoes & Jewelry · Home & Kitchen · Office Products · Tools & Home Improvement · Health & Household · Patio, Lawn & Garden · Electronics · Cell Phones & Accessories · Video Games · Grocery & Gourmet Food


One Stop Market Product Showcases




Pre-baked Gingerbread House Kit Value Pack, 17 oz., Pack of 2, Total 34 oz.
★ ★ ★ ★ ★ 1 Review
\$19.99
[Add to Cart](#)




V8 +Energy, Healthy Energy Drink, Steady Energy from Black and Green Tea, Pomegranate Blueberry, 8 Ounce Can, Pack of 24
★ ★ ★ ★ ★ 12 Reviews
\$14.47
[Add to Cart](#)




Elmwood Inn Fine Teas, Orange Vanilla Caffeine-free Fruit Infusion, 16-Ounce Pouch
★ ★ ★ ★ ★ 4 Reviews
\$19.36
[Add to Cart](#)




Belle Of The Ball Princess Sprinkle Mix | Wedding Colorful Sprinkles | Cake Cupcake Cookie Sprinkles | Ice cream Candy Sprinkles | Yellow Gold Red Royal Red Rose Icing Flowers Decorating Sprinkles, 9OZ
★ ★ ★ ★ ★ 12 Reviews
\$23.50
[Add to Cart](#)




So Delicious Dairy Free CocoWhip Light, Vegan, Non-GMO Project Verified, 9 oz. Tub
★ ★ ★ ★ ★ 12 Reviews
\$15.62
[Add to Cart](#)




Cheongun Sweet Potato Starch Powder 500g, 2ea (Sweet Potato 55%, Corn 45%)
\$34.00
[Add to Cart](#)




Q Mixers Premium Ginger Ale: Real Ingredients & Less Sweet, 6.7 Fl Oz (24 Bottles)
★ ★ ★ ★ ★ 12 Reviews
\$68.50
[Add to Cart](#)




Stove Top Turkey Stuffing Mix (12 oz Boxes, Pack of 2)
★ ★ ★ ★ ★ 12 Reviews
\$8.49
[Add to Cart](#)




Briss DME - Pilsen Light - 1 lb Bag
\$12.99
[Add to Cart](#)



Tony Chachere's More Spice Creole Seasoning - 14 oz
★ ★ ★ ★ ★ 12 Reviews
\$7.92
[Add to Cart](#)



Crunchy Rice Rollers - Gluten Free - Vegan - 3.5 oz Individual Packs (4 Packs of 8 Rollers)
★ ★ ★ ★ ★ 12 Reviews
\$11.50
[Add to Cart](#)



Lobster Cobbler Pot Pie - Gourmet Frozen Seafood Appetizers (Set of 8 Trays)
\$411.76
[Add to Cart](#)

Items 1 to 12 of 24 total
1 2 >

Privacy and Cookie Policy
Search Terms
Advanced Search
Orders and Returns
Contact Us

Copyright © 2013-present Magento, Inc. All rights reserved.
Help Us Keep Magento Healthy Report All Bugs

QUESTION: Which item is an energy drink?
(Answer - B)

A) Q Mixers Premium Ginger Ale: Real Ingredients & Less Sweet, 6.7 Fl Oz (24 Bottles)
B) V8 +Energy, Healthy Energy Drink, Steady Energy from Black and Green Tea, Pomegranate Blueberry, 8 Ounce Can, Pack of 24
 C) Elmwood Inn Fine Teas, Orange Vanilla Caffeine-free Fruit Infusion, 16-Ounce Pouch
 D) So Delicious Dairy Free CocoWhip Light, Vegan, Non-GMO Project Verified, 9 oz. Tub
 E) Briss DME - Pilsen Light - 1 lb Bag
 F) Cheongun Sweet Potato Starch Powder 500g, 2ea (Sweet Potato 55%, Corn 45%)

QUESTION: Which product is the most expensive on the page? (Answer - C)

A) Pre-baked Gingerbread House Kit Value Pack, 17 oz., Pack of 2, Total 34 oz.
 B) Q Mixers Premium Ginger Ale: Real Ingredients & Less Sweet, 6.7 Fl Oz (24 Bottles)
C) Lobster Cobbler Pot Pie - Gourmet Frozen Seafood Appetizers (Set of 8 Trays)
 D) Tony Chachere's More Spice Creole Seasoning - 14 oz
 E) Stove Top Turkey Stuffing Mix (12 oz Boxes, Pack of 2)
 F) So Delicious Dairy Free CocoWhip Light, Vegan, Non-GMO Project Verified, 9 oz. Tub

QUESTION : Which product is a tea infusion sold in a 16-ounce pouch? (Answer - A)

A) Elmwood Inn Fine Teas, Orange Vanilla Caffeine-free Fruit Infusion, 16-Ounce Pouch
 B) V8 +Energy, Healthy Energy Drink, Steady Energy from Black and Green Tea, Pomegranate Blueberry, 8 Ounce Can, Pack of 24
 C) So Delicious Dairy Free CocoWhip Light, Vegan, Non-GMO Project Verified, 9 oz. Tub
 D) Crunchy Rice Rollers - Gluten Free - Vegan - 3.5 oz Individual Packs (4 Packs of 8 Rollers)
 E) Q Mixers Premium Ginger Ale: Real Ingredients & Less Sweet, 6.7 Fl Oz (24 Bottles)
 F) Tony Chachere's More Spice Creole Seasoning - 14 oz

QUESTION : Which product shows the highest customer rating percentage? (Answer - B)

A) V8 +Energy, Healthy Energy Drink, Steady Energy from Black and Green Tea, Pomegranate Blueberry, 8 Ounce Can, Pack of 24
B) Elmwood Inn Fine Teas, Orange Vanilla Caffeine-free Fruit Infusion, 16-Ounce Pouch
 C) Q Mixers Premium Ginger Ale: Real Ingredients & Less Sweet, 6.7 Fl Oz (24 Bottles)
 D) Stove Top Turkey Stuffing Mix (12 oz Boxes, Pack of 2)
 E) Crunchy Rice Rollers - Gluten Free - Vegan - 3.5 oz Individual Packs (4 Packs of 8 Rollers)
 F) Tony Chachere's More Spice Creole Seasoning - 14 oz

Figure 1. Examples from the repropoed VWA multiple-choice dataset constructed on shopping webpages. For each webpage, we generate 10 questions with 6 answer choices, among which 1 is correct. The questions cover diverse reasoning types, including product category recognition, price comparison, customer rating analysis, and attribute identification.

My Account My Wish List Sign In Welcome to One Stop Market Create an Account

One Stop Market Search entire store here...

Advanced Search

Beauty & Personal Care Sports & Outdoors Clothing, Shoes & Jewelry Home & Kitchen Office Products Tools & Home Improvement

Health & Household Patio, Lawn & Garden Electronics Cell Phones & Accessories Video Games Grocery & Gourmet Food

Home > Beauty & Personal Care > Fragrance > Men's

Men's

Shop By

Shopping Options

Price

\$0.00 - \$99.99 (238)

\$100.00 - \$199.99 (20)

\$200.00 - \$299.99 (4)

\$300.00 and above(2)

Compare Products
You have no items to compare.

My Wish List
You have no items in your wish list.

Sort By: Position

Items 1-12 of 265

Bever Hills Polo Club Classic by Beverly Hills Polo Club
Fragrances Eau De Toilette Spray, 3.4 oz
\$30.95

Add to Cart

Territoire Eau De Parfum, Men's Cologne (Platinum)
\$13.40

Add to Cart

Imaginary Men's
Trespasser Eau de Parfum Unisex Spray (1.7 Fl Oz)
\$95.00

Add to Cart

Women's Patented Male Pheromone Fragrance "CASA" Attraction Perfum
\$26.97

Add to Cart

Trussardi Uomo The Red By Trussardi for Men - 3 Pc Gift Set
3.4oz Edt Spray, 3.4oz Shampoo Anad Shower Gel, Washbag, 3count
\$51.18

Add to Cart

Hermes Terre D'Ermenegildo Zegna
Perfume 2 Pc Set FOR MEN (4.2 Oz Eau De Toilette + 1.0 Oz Eau De Parfum Refillable Spr)
\$124.99

Add to Cart

Club, 6 oz Body Spray for Men
★★★★★ 12 Reviews
\$14.37

Add to Cart

★★★★★ 12 Reviews
\$18.68

Add to Cart

1 2 3 4 5 >

Show 12 per page

Privacy and Cookie Policy

Search Terms

Advanced Search

Orders and Returns

Contact Us

Copyright © 2013-present Magento, Inc. All rights reserved.

Help Us Keep Magento Healthy Report All Bugs

Enter your email address

Subscribe

QUESTION : What number of items per page is currently selected?
(Answer - A)

A) 12 B) 24
C) 36 D) 1
E) 2 F) 5

QUESTION : According to the toolbar, how many total items are in this category?
(Answer - A)

A) 265 B) 12
C) 238 D) 20
E) 4 F) 2

QUESTION : Which view mode is currently active on the page?
(Answer - A)

A) Grid B) List
C) Position D) Price
E) Men's F) My Cart

QUESTION : Which Sort By option is selected by default?
(Answer - A)

A) Position B) Product Name
C) Price D) Grid
E) List F) Next

QUESTION : From the Shop By price filter, how many items are in the \$100.00 - \$199.99 range?
(Answer - B)

A) 238 B) 20
C) 4 D) 2
E) 12 F) 265

QUESTION : In the pagination, what is the last page number shown?
(Answer - E)

A) 1 B) 2
C) 3 D) 4
E) 5 F) Next

Figure 2. Examples from the repropoed VWA multiple-choice dataset constructed on shopping webpages.

Dataset	Prompt Template
<p>DocVQA & InfoVQA</p>	<p>BACKWARD QUERY GENERATION</p> <p>System Message: You are a question generator. Given an observation of a document (infographic) and the correct Answer, produce a single, concise, unambiguous question whose answer is exactly the given Answer and should be asking about the information provided by the observation; when grounded ONLY on the observation provided. Rules: (1) Return ONLY the question text. (2) Avoid ambiguous wording; ensure a 1-to-1 mapping, this means the query should not have answers to it other than the correct Answer provided to you. (3) No extra commentary, quotes, or prefixes. (4) Make sure that the question is asking about the observed document (infographic), don't propose random answers to fit the answer.</p> <p>User Message: Base rule: You need to generate a question. Here is the structure of the examples: <i>few shot examples:</i></p> <p>Example i OBS: [TEXT_OBS] or [IMAGE_OBS] Answer: [ANS] Given the observation of the document (infographic), I will come up with a document (infographic) facts based Query that can be answered by the given answer. Query is: [QUESTION]</p> <p><i>The remaining examples follow the same structure.</i></p> <p>Visual focus: In this task, the OBS is an image version of the document (infographic). You must generate the question based solely on visible information in the image, without assuming or inferring any unseen text or external knowledge. Textual focus: In this task, the OBS is ocr and captioning of an document (infographic) image. OBS: [TEXT_OBS] or [IMAGE_OBS] Answer: [ANS] Given the observation of the document (infographic), I will come up with a document (infographic) facts based Query that can be answered by the given answer. Query is:</p> <p>Assistant Message: (Generated response)</p> <hr/> <p>CYCLE VERIFICATION</p> <p>User Message: You are a helpful assistant for document VQA. Answer with the exact final answer only. OBS: [TEXT_OBS] or [IMAGE_OBS] Query: [QUESTION] Answer concisely with only the final answer.</p> <p>Assistant Message: (Generated response)</p>

Table A2. **Prompt templates for each dataset and prompt type.** The BACKWARD QUERY GENERATION prompt infers plausible multimodal queries conditioned on each candidate answer; and the CYCLE VERIFICATION prompt assesses whether forward inference from these queries reconstructs the original answer, providing the cycle-consistency supervision signal.

Dataset	Prompt Template
ChartQA	<p>BACKWARD QUERY GENERATION</p> <p>Image-specific part: [IMAGE_OBS] You are writing ONE short, clear, self-contained question about a chart image such that the answer equals a GIVEN TARGET ANSWER.</p> <p>Text-specific part: You are writing ONE short, clear, self-contained question about a chart based ONLY on the following caption text. The question must have the GIVEN TARGET ANSWER. [TEXT_OBS]:</p> <p>Shared rules and ending: Rules:</p> <ul style="list-style-type: none"> - Must be answerable using chart visual information only. - Must not reveal or restate the answer. - Avoid yes/no, multiple choice, multi-part questions. - Include necessary qualifiers (series, category, unit, timestamp). - Length: 1–2 sentences, no explanation. <p>Target Answer: [ANSWER]. Write the question now. Return only the question string.</p>
	<p>CYCLE VERIFICATION</p> <p>Image-based QA Prompt: [IMAGE_OBS] You are given a chart image and a question. Use ONLY the information that is explicitly visible in the chart (titles, labels, legends, tick marks, data labels, notes). Question: [QUESTION] Answer concisely with plain or numeric text only (no reasoning, no steps, no formatting).</p> <p>Text-based QA Prompt: You are given a chart description in JSON extracted from an image. Answer the user’s question using ONLY the information contained in the JSON. JSON: [TEXT_OBS] Question: [QUESTION] Answer concisely with plain or numeric text only (no reasoning, no steps, no formatting).</p>

Table A3. **Prompt templates for each dataset and prompt type.** The BACKWARD QUERY GENERATION prompt infers plausible multimodal queries conditioned on each candidate answer; and the CYCLE VERIFICATION prompt assesses whether forward inference from these queries reconstructs the original answer, providing the cycle-consistency supervision signal.