54

55

# Search-in-the-Chain: Interactively Enhancing Large Language Models with Search for Knowledge-intensive Tasks

Anonymous Author(s)

# ABSTRACT

Making the contents generated by Large Language Model (LLM) such as ChatGPT, accurate, credible and traceable is crucial, especially in complex knowledge-intensive tasks that require multi-step reasoning and each step needs knowledge to solve. Incorporating Information Retrieval (IR) to provide LLM with external knowledge is good potential to solve this problem. However, where and how to introduce IR into LLM is a big challenge. Previous work has the problems that wrong knowledge retrieved by IR will mislead the LLM and interaction between IR and LLM breaks the reasoning chain of LLM. In this paper, we propose a novel framework named Search-in-the-Chain (SearChain) for the interaction between LLM and IR to solve the challenges. First, LLM generates the global reasoning chain named Chain-of-Query (CoQ) where each node consists of an IR-oriented query and the answer to the query. Second, IR verifies the answer of each node of CoQ. It corrects the answer that is not consistent with the retrieved information when IR gives high confidence, which improves the credibility. Third, LLM can indicate its missing knowledge in CoQ and rely on IR to provide this knowledge to LLM. These three operations improve the accuracy of LLM for complex knowledge-intensive tasks in terms of reasoning ability and knowledge. Finally, SearChain generates the reasoning process and marks references to supporting documents for each reasoning step, which improves traceability. Interaction with IR in SearChain forms a novel reasoning path: node-identify Depth-first Search on a tree, which enables LLM to dynamically modify the direction of reasoning. Experiments show that SearChain outperforms recent state-of-the-art baselines on complex knowledge-intensive tasks including multi-hop Q&A, slot filling, fact checking, and long-form Q&A.

# **CCS CONCEPTS**

Computing methodologies → Natural language processing.

# KEYWORDS

Retrieval-augmented model, Large Language Models

## ACM Reference Format:

Anonymous Author(s). 2024. Search-in-the-Chain: Interactively Enhancing Large Language Models with Search for Knowledge-intensive Tasks. In Proceedings of the ACM Web Conference 2024 (www '24), May 13–17, 2024, Singapore. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/ nnnnnnn.nnnnnn

WWW '24, May 13–17, 2024, Singapore

57 https://doi.org/10.1145/nnnnnnnnn
 58

# 1 INTRODUCTION

Large Language Models (LLMs) such as ChatGPT have shown promising performance in various natural language processing tasks [2, 41]. As LLMs are gradually used in various fields, we need to pay attention to how to ensure that the contents generated by LLMs are accurate, credible and traceable, especially in the complex knowledge-intensive tasks that require multi-step reasoning and each step needs corresponding knowledge to solve [20, 37]. Many studies have shown that LLMs have trouble in: (1) compositional reasoning over multiple knowledge [21], (2) memorization of longtail and real-time knowledge [12] and (3) avoiding hallucination that is inconsistent with the facts [1], which affects the accuracy and credibility of LLMs for complex knowledge-intensive tasks. Besides, context-only generation without any supporting evidence causes less traceability and makes people less trust in the LLM-generated content. Retrieval-augmented method has good potential to solve these problems because it combines the knowledge of the model with external knowledge bases [9, 11, 17].

59 60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

However, how to introduce IR into LLM is not a trivial thing. There are three main challenges. C-1: Directly inserting IR into the reasoning process of LLM such as Self-Ask [21], LTM [42], React [36] and DSP [14] leads to breaking the reasoning chain of LLM. Because in these methods, LLM can only reason a local subquestion in each generation. Although AgentGPT and PS [31] first plan sub-questions and then solve them, they are not suitable for scenarios where the next sub-question is dependent on the answer of the previous sub-questions, which is common for complex tasks, such as the multi-hop QA. C-2: When there is a conflict in the knowledge of IR and LLM, for the knowledge that the LLM has correctly memorized, it risks being misled by IR if IR retrieves the wrong information. Therefore, it is important to decouple the knowledge of LLM and IR to make sure that IR only provides the knowledge that LLM really needs to avoid the misleading of LLM by IR, which is not considered in previous methods. C-3: Previous methods cannot dynamically modify the reasoning direction.

In this paper, we propose a novel framework named Searchin-the-Chain (SearChain) to effectively combine LLM with IR to solve the above challenges (Figure 1). SearChain and previous methods both need multiple IR-LLM interaction rounds, but the former works at the chain level, while the latter only deals with a node. In each round, SearChain performs reasoning, verification, and completion. After the interaction, SearChain performs tracing to generate the final content. Specifically, in each round, first, LLM exploits in-context learning to construct a Chain-of-Query (CoQ), which is a reasoning chain to decompose and solve complex questions. Each node of the chain consists of an IR-oriented query, the answer generated by LLM for this query and a flag indicating whether LLM needs additional knowledge. Different from previous methods in which LLM can only perform one-step reasoning (only a node) when interacting with IR, CoQ is a complete chain. This design

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

<sup>© 2024</sup> Copyright held by the owner/author(s).

ACM ISBN 978-x-xxxx-x/YY/MM.

135

136 137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174



Figure 1: Interaction between IR and LLM in SearChain. First, SearChain makes LLM plan a CoQ where each node is a queryanswer pair. Then, IR interacts with each node of CoQ to perform verification and completion. If IR detects that a node needs to be corrected or provided with knowledge, it gives feedback to LLM and LLM re-generates a new CoQ, which is the new branch of the tree. This process is the node-identify Depth-first Search on a tree called Tree-of-Reasoning (correct reasoning path is green). Final content includes the reasoning process and references to supporting documents.

avoids IR from breaking the reasoning chain (C-1). Second, IR interacts with each node of CoQ to perform verification and completion. In verification, IR verifies the answer of each node. In case when the LLM-generated answer is not consistent with the retrieved information and IR gives high confidence, IR gives feedback to LLM to help it correct the answer and re-generate the correct CoQ. In completion, IR determines whether the node has missing knowledge from the flag of the node and provides this knowledge to LLM to help it re-generate CoQ. LLM gradually generates the correct CoQ through multiple rounds of interaction with IR. The above design provides LLM with the knowledge it really needs to alleviate the misleading caused by IR to LLM (C-2), which improves accuracy. IR verifies and corrects the knowledge in the reasoning process of LLM based on external knowledge bases, which improves credibility. After the interaction, SearChain performs tracing to generate the reasoning process and marks references to supporting documents for each reasoning step, which is used as the final content returned to the user. This improves the traceability of knowledge in the generated contents. Interaction with IR in SearChain transforms the reasoning path from a chain to node-identify Depth-first Search on a tree called Tree-of-Reasoning (ToR). CoQ generation can be seen as a part of Depth-first Search and IR can identify the nodes that need more information (C-3). This enables LLM to dynamically modify the reasoning direction. This paper's main contributions are:

(1) We highlight the challenges in introducing IR into LLM from the perspectives of reasoning and knowledge.

(2) We propose a novel framework called SearChain to effectively combine LLM and IR. SearChain not only stimulates the knowledgereasoning ability of LLM but also uses IR to identify and offer the knowledge that LLM really needs, which helps to improve accuracy and credibility. Moreover, SearChain can also mark references to supporting documents for the knowledge involved in the generated contents, which helps to improve the traceability of the knowledge. (3) Interaction with IR in SearChain forms a novel reasoning path: node-identify Depth-first Search on a tree, which enables LLM to dynamically modify the direction of reasoning.

(4) Experiment shows that SearChain outperforms state-of-theart baselines on complex knowledge-intensive tasks including multihop question-answering, slot filling, fact checking and long-form question-answering. Code will be released at https://github.com.

# 2 RELATED WORK

# 2.1 Chain-of-Thought Prompting

Chain-of-thought [33] proposes the method that uses few-shot examples to enable LLM to give intermediate reasoning results when solving complex problems and improves the reasoning ability. [15] uses "Let's do it step by step" as prompt to achieve promising zeroshot performance. Auto-CoT exploits language models to automatically construct few-shot learning examples for CoT [39]. There are also many studies that cover other aspects of CoT such as selfconsistency [32], usage of small and medium size models [38] and selection [7]. Besides, there are studies that iteratively use LLM to decompose complex questions and answer sub-questions step by step. These methods include Least-to-Most [42], Dynamic Leastto-Most [4], Self-Ask [21] and DSP [14]. Chain-of-Query of our method is also inspired by CoT. However, previous studies focus on giving intermediate reasoning results or decomposing complex questions and answering sub-questions step by step. They focus on how to solve local sub-questions while ignoring the global planning of the reasoning chain. Although AgentGPT and PS [31] first plan each sub-question and then solve them, they are not suitable for scenarios where the next sub-question needs the answer of the previous sub-questions to generate, which is common for complex knowledge-intensive tasks (multi-hop QA). CoQ of our method makes LLM construct a global reasoning chain where each node is

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

a query-answer pair. This design not only improves the knowledgereasoning ability but also provides the interface for IR to be deeply involved in the reasoning process of LLM.

#### **Retrieval-augmented Language Models** 2.2

Many studies have shown that retrieval-augmented methods can connect language models with external corpus to get promising performance in various natural language tasks such as open-domain question answering [9, 11, 17], language modeling [3, 19] and enhancing the factuality [22]. Recently, some studies enable LLM to interact with IR via in-context learning [14, 21, 25, 36]. In these methods, the interaction between IR and LLM makes the reasoning of LLM not continuous. LLM can only perform one-step reasoning at each inference. Our method makes LLM generate a global reasoning chain called Chain-of-Query at each inference, the hidden states in generation are used as the sequential dependency, which introduces stronger logical relationship between each reasoning step. Besides, previous methods can only provide information to the LLM but cannot assist LLM in correcting erroneous information or avoid the negative effect of IR on LLM, which makes the reasoning of LLM still in a one-dimensional chain. Our method makes IR interact with each node of the chain. IR only provides LLM with its missing knowledge and corrects the answers that are not consistent with the retrieved information when IR is confident enough. This design mitigates the negative effect of IR on LLM and transforms the reasoning path from chain to node-identify Depth First Search on a tree to enable LLM to modify the reasoning direction.

#### **OUR METHOD** 3

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

This section introduces the design of Search-in-the-Chain (SearChain). In SearChain, IR and LLM conduct multiple rounds of interaction. In each round, first, LLM acts as the commander to plan the global reasoning chain for the complex input questions called Chain-of-Query (CoQ). Each node of the CoQ consists of an IR-oriented query, the answer generated by LLM for this query and a flag indicating whether LLM needs additional knowledge. Then, IR interacts with each node of CoQ and performs the completion and verification to provide LLM with missing knowledge and correct the wrong answers. LLM re-generates new CoQ based on feedback from IR. Multiple rounds of interaction help LLM to gradually generate the correct CoQ according to the external knowledge base, which improves accuracy and credibility. Finally, SearChain performs tracing to generate the whole reasoning process and marks references to supporting documents for each reasoning step, which is used as the final content returned to the user. This improves the traceability of generated content. Interaction with IR in SearChain transforms the reasoning path from a chain to node-identify Depth-first Search on a tree called Tree-of-Reasoning (ToR), which enables LLM to dynamically modify the reasoning direction. Besides, SearChain decouples the knowledge of LLM and IR and provides LLM with the knowledge it really needs to alleviate the misleading of LLM.

## 3.1 Comparison with Previous Methods

Figure 2 shows the difference between our method and previous retrieval-augmented methods (Self-Ask [21], React [36], DSP [14], etc.) in solving complex knowledge-intensive questions.



Figure 2: Comparison with previous methods.

(1) Local vs. Global. For a complex question that needs multistep reasoning, previous methods directly insert IR into the multistep reasoning process, causing LLM can only reason a local subquestion such as node (A) in each generation. This breaks the reasoning chain of LLM. Our method proposes Chain-of-Query to provide the interactive interface for IR on the premise of ensuring the coherence of reasoning chain (plan a global chain for question *Q* such as  $\mathbb{A} \to \mathbb{B} \to \mathbb{C} \to \mathbb{D}$  in each generation). (solves *C*-1) (2) Directly Provide vs. Verify and Complete. Previous methods directly provide the retrieved information to the LLM. When the retrieved information is incorrect, the LLM runs the risk of being misled. In our method, IR only corrects inconsistent information in Chain-of-Query when IR is confident enough, and provides the information that LLM does not know via flags on Chain-of-Query, which mitigates the negative effect of IR on LLM. (solves C-2)

(3) Chain vs. Tree. Previous methods cannot modify the reasoning direction in time as necessary. Our method transforms the reasoning path from a chain to node-identify Depth-first Search on a tree by introducing the verification and completion from IR, which enables LLM to dynamically modify the direction of reasoning. (solves C-3)

# 3.2 Chain-of-Query Generation

In SearChain, we use in-context learning [33] to prompt large language model to construct a global reasoning chain for complex question Q named Chain-of-Query (CoQ):

$$\operatorname{CoQ} = (q_1, a_1) \to (q_2, a_2) \to \dots \to (q_n, a_n), \tag{1}$$

which is the branch of Tree-of-Reasoning. Each node  $(q_i, a_i)$  of CoQ consists of an IR-oriented query  $q_i$  and its answer  $a_i$ .  $q_1$  ...  $q_n$  are the sub-questions that need to be solved in the reasoning process of solving Q. CoQ generation is applied to each round of interaction between LLM and IR. In the first round, the prompt used to make LLM generate CoQ is shown in Figure 3. The prompt

349	Construct a global reasoning chain for this complex [Question] : Where do
350	Greyhound buses that are in the birthplace of Spirit If 's performer leave from?
351	You should generate a query to the search engine based on what you already know at each step of the reasoning chain starting with [Ouerv]
352	If you know the answer for [Ouery] generate it starting with [Answer]
353	You can try to generate the final answer for the [Question] by referring to the
354	[Query]-[Answer] pairs, starting with [Final Content].
355	If you don't know the answer, generate a query to search engine based on what you already know and do not know, starting with [Unsolved Query].
356	For example:
357	[Question]: Where do Greyhound buses that are in the birthplace of Spirit If 's
358	performer leave from?
550	[Query 1]: Who is the performer of Spirit If?
359	If you don't know the answer:
360	[Unsolved Query]: Who is the performer of Spirit If?
361	If you know the answer:
262	[Answer 1]: The performer of Spirit If is Kevin Drew.
302	[Query 2]: Where was Kevin Drew born?
363	If you don't know the answer:
364	[Unsolved Query]: Where was Kevin Drew born?
365	If you know the answer:
044	[Answer 2]: Toronto.
300	[Query 3]: Where do greyhound buses in Toronto leave from?
367	If you don't know the answer:
368	[Unsolved Query]: Where do greyhound buses in Toronto leave from? - Node
369	If you know the answer:
507	[Answer 3]: Toronto Coach Terminal.
370	[Final Content]: The performer of Spirit If is Kevin Drew [1]. Kevin Drew
371	was born in Toronto [2]. Greyhound buses in Toronto leave from Toronto
372	Coach Terminal [3]. So the final answer is Toronto Coach Terminal.

Figure 3: Prompt to make LLM generate Chain-of-Query.

starts with "Construct a global reasoning chain" to make LLM know that the main task is to generate a global reasoning chain in each generation. "Global" means that LLM needs to plan a complete reasoning chain for the complex question, rather than answer the question directly or only solve "local" sub-questions (comparison shown in Figure 2). At each node of the chain, LLM focuses on generating the IR-oriented query and gives the answer if LLM knows. If LLM does not know the answer, it should mark the query with "[Unsolved Query]", which is a flag indicating the missing of knowledge. In subsequent rounds, when a node needs IR to correct or provide missing knowledge, LLM generates a new CoQ according to the feedback of IR to dynamically modify the reasoning direction. The design for this scenario will be introduced in Section 3.3. The generation of CoQ is a complete Depth-first Search for Q, which avoids IR from breaking the reasoning chain of LLM. Experiments (Section 4.3.3) also show that for the difficult sub-question, CoQ enables LLM to solve it by more reasoning steps such as rewriting or further decomposing the sub-question while baselines tend to stop reasoning. It is because baselines focus on solving current local sub-questions while ignoring the global planning of the reasoning chain. The global perspective in CoQ makes LLM try harder to explore possible answers when facing intermediate difficulties. 

# 3.3 Interaction with Information Retrieval

In each round of interaction, LLM passes the generated CoQ to IR. IR verifies and completes the information for each node  $(q_i, a_i)$ of CoQ and feeds back to LLM to help it generate more correct CoQ as the new branch of ToR (Tree-of-Reasoning). Besides, IR records the corresponding retrieved documents for each node of 

Algorithm 1: Description of the Interaction with IR.	
<b>Initialize :</b> Processed queries: <i>M</i> = <i>null</i> ;	
Correct reasoning path: $R = null$ ;	
Interaction rounds: $r = 0$ ;	
Feedback: $F = null$ ; ToR: $T = Q$ ;	
<b>Function</b> $\underline{IR}(q_i, a_i)$ :	
$d_i = \text{Retrieval}(\underline{q_i}); // \text{Retrieve Top-1 document } d_i \text{ for } q_i.$	
$g, f = \operatorname{Reader}(q_i, d_i);$	
// Extract answer g from $d_i$ and give confidence f.	
if <u>q<sub>i</sub> is Unsolved Query</u> then	
// Completion.	
R.add $(q_i, g, d_i)$ ; // Record the correct node.	
<b>return</b> PromptForComplete( $\underline{q_i, g, d_i}$ );	
if $f > \theta$ and NotEqual $(g, a_i)$ then	
// Verification.	
$R.add(q_i, g, d_i);$ // Record the correct node.	
<b>return</b> PromptForVerify( $\underline{q_i, g, d_i}$ );	
<i>R</i> .add $(q_i, a_i, d_i)$ ; <b>return</b> "Pass";	
Function Traverse(CoQ):	
foreach $(q_i, a_i)$ in $CoQ$ do	
if not DuplicateQuery $(q_i, M)$ then	
// If $q_i$ has not been processed.	
$F = \operatorname{IR}(\underline{q_i, a_i}); M.\operatorname{add}(q_i);$	
<b>if</b> <u>not <math>F ==</math> "Pass"</u> <b>then return</b> $F$ ;	
return "Finish" ·	
Function Main( $\underline{Q},F$ ):	
while not (F == "Finish" or $r > r_{max}$ ) do	
CoQ = ChainGenerate(Q, F);	
// LLM generate the new Chain-of-Query CoQ.	
I.AddCh11d( $\underline{CoQ}$ ); // Add the new branch to T.	
$F = \text{Traverse}(\underline{CoQ}); // Interact with IR.$	
r = r + 1; // Update the number of interaction rounds r.	
return Tracing(T, R)	

CoQ as its supporting documents, which enhances the traceability of LLM-generated content. The description of interaction is shown in Algorithm 1. IR interacts with each node  $(q_i, a_i)$  of CoQ, retrieves the Top-1 document  $d_i$  for  $q_i$  as the supporting document, and judges whether to verify or complete it according to the type of  $q_i$ . When all the queries of CoQ do not need to be corrected or completed, or the maximum number of interaction rounds is reached, the interaction ends. SearChain traces back the correct reasoning path of ToR and refers to each node of the path to generate the final content with marked references to supporting documents for knowledge of each node.

Verification. Verification aims to guarantee the correctness of  $a_i$  in each node  $(q_i, a_i)$  of CoQ based on the external knowledge base, which improves the accuracy and credibility of generated content. Specifically, given the retrieved Top-1 document  $d_i$  for  $q_i$ , a Reader [13] that has been trained on open-domain QA datasets [13] is used to extract the answer g for  $q_i$  from  $d_i$  with its confidence f(*f* is a predicted value that measures whether *q* can answer  $q_i$ ):

 $s = \arg \max(\operatorname{softmax}(\mathbf{Hw}_s)), e = \arg \max(\operatorname{softmax}(\mathbf{Hw}_e)),$ 

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

$$g = d_i[s:e], f = \mathbf{H}_{[CLS]}\mathbf{w}_f, (\mathbf{w}_s, \mathbf{w}_t, \mathbf{w}_f \in \mathbb{R}^E),$$

465

466

where  $\mathbf{H} \in \mathbb{R}^{L \times E}$  is the sequence of last hidden states for the input 467 text "[CLS] $q_i$ [SEP] $d_i$ ", *L* is the length and *E* is hidden dimension. 468 H<sub>[CLS]</sub> is the last hidden state of [CLS] token. Then, SearChain 469 judges whether the answer  $a_i$  given by LLM is consistent with 470 the retrieved information according to (1) whether q appears in 471  $a_i$  (for short-form generation tasks such as multi-hop QA and slot 472 473 filling) or (2) whether ROUGE [18] between  $a_i$  and  $d_i$  is greater 474 than the threshold  $\alpha$  (for long and free-form generation tasks such as ELI5 [6]). If  $a_i$  is not consistent with retrieved information and 475 Reader is confident enough ( $f > \theta$ ,  $\theta$  is a threshold to alleviate the 476 negative effect of IR on LLM), a prompt is constructed to help LLM 477 correct the answer *a<sub>i</sub>*. The template of the prompt is: "According 478 to the Reference, the answer for  $q_i$  should be  $g_i$ , you can change your 479 480 answer and continue constructing the reasoning chain for [Question]: Q. Reference:  $d_i$ .". This round is over. LLM receives the feedback of 481 IR, gives the new answer  $a'_i$  for q, and generates a new CoQ with 482 483  $(q_i, a'_i)$  as the root node, which is the new branch of ToR.

Completion. Completion aims to provide LLM with missing 484 knowledge in nodes of CoQ, which improves the accuracy of gen-485 486 erated contents. Specifically, in CoQ generation (Section 3.2), LLM 487 marks "[Unsolved Query]" for the unsolvable query. For the unsolvable query  $q_i^*$ , IR extracts the answer  $q^*$  from retrieved document 488  $d_i^*$  as described in Verification. Regardless of whether f is greater 489 than the threshold  $\theta$ ,  $g^*$  and  $d_i^*$  will be fed back to the LLM in the 490 form of a prompt because the LLM cannot solve  $q_i^*$ . The template of 491 the prompt is: "According to the Reference, the answer for  $q_i^*$  should 492 493 be  $q^*$ , you can give your answer and continue constructing the reasoning chain for [Question]: Q. Reference:  $d_i^*$ .". This round is over. 494 LLM receives the feedback, gives the answer  $a_i^*$  to solve the query 495  $q_i^\ast$  and generates a new CoQ with  $(q_i^\ast, a_i^\ast)$  as the root node, which 496 497 is the new branch of ToR.

Tracing. Tracing aims to generate the reasoning process and 498 499 mark references to supporting documents for each reasoning step, 500 which is used as the final content returned to the user. This improves the traceability of each knowledge in the generated content. 501 502 Specifically, SearChain records the documents retrieved for each node on the correct reasoning path of Tree-of-Reasoning as the 503 supporting documents. SearChain prompts LLM to generate the 504 final content by referring to nodes on the correct path and mark 505 506 references to the supporting documents for the corresponding subfragments of the generated content (final content of Figure 1). The 507 prompt is "You can try to generate the final answer for the [Ques-508 509 tion] by referring to the [Query]-[Answer] pairs, starting with [Final 510 Content]. [Query 1]:  $q_1$ . [Answer 1]:  $a_1 \dots$  [Query m]:  $q_m$ . [Answer m]: 511  $a_m$ .". This design enables the user to acquire the related documents 512 of the knowledge involved in each step of reasoning. We believe 513 that it is a promising task to mark references to supporting documents on sub-fragments of complex content generated by LLM. 514 Our approach provides a novel and effective approach to solve this 515 task by retrieving supporting documents for each sub-questions 516 involved in the reasoning process of LLM without any supervised 517 data (texts with citation annotations) and training of the LLM. 518

Node-identify Depth-first Search. Compared with previous
 retrieval-augmented methods, interaction with IR in SearChain
 forms a novel reasoning path: node-identify Depth-first Search

on a tree. In each generation, LLM generates a CoQ to perform continuous reasoning on complex questions until the final answer is generated or an unsolvable sub-question is encountered. This can be seen as a part of Depth-first Search (DFS). However, different from traditional DFS algorithm [28], "node-identify" in SearChain means that when a search in one direction is terminated, SearChain does not return to its parent node, but dynamically identifies the node that needs to be corrected or completed via verification and completion in IR and re-generates a new CoQ started with this node. The interaction process between IR and LLM in SearChain is the process of constructing a tree using node-identify DFS, which enables LLM to dynamically modify the reasoning direction.

# **4 EXPERIMENTS**

In this section, we compare SearChain with recent related baselines on complex knowledge-intensive tasks and conduct the analysis.

# 4.1 Experimental Setup

4.1.1 **Datasets and Evaluation Metric**. We select four classic complex knowledge-intensive tasks including multi-hop questionanswering (HotpotQA (HoPo) [34], Musique (MQ) [30], WikiMulti-HopQA (WQA) [10] and StrategyQA (SQA) [8]), slot filling (zsRE [16], T-REx [5]), fact checking (FEVER [29]) and long-form questionanswering (ELI5 [6]). These tasks require LLM to perform multistep reasoning on complex questions, and each step requires corresponding knowledge to solve. As for the evaluation metrics, for ELI5 whose ground truth is long and free-form, we use ROUGE-L [18] as the metric. For other tasks, we use whether the ground truth answer is contained within the generated answer (i.e, cover-EM [23]) as the metric. Following DSP [14] and Self-Ask [21], we evaluate the model on full development datasets of MQ and HoPo, BIG-bench [26] datasets on SQA and subsets of WQA, zsRE, T-REx, FEVER and ELI5 (each subset has 1.2k questions).

4.1.2 **Baselines.** Our baselines can be divided into two categories, one is about improving the reasoning ability of LLM on complex tasks (CoT [33], CoT-SC [32], Auto-CoT [39], Recite-and-answer [27] and Least-to-Most [42]), and the other is not only introducing IR to LLM but also improving the reasoning ability (Direct<sup>1</sup>, Self-Ask [21], ToolFormer<sup>2</sup> [25], React [36], DSP [14], Verify-and-Edit (combined with CoT-SC) [40] and Tree-of-Thought [35]). AgentGPT and PS [31] use Plan-and-Solve paradigm, we also reproduce this as one of the baselines.

4.1.3 **Implementation**. The large language model we used is *gpt*-3.5-turbo provided from API of OpenAI<sup>3</sup> and the retrieval model we used is ColBERTv2 [24] (following DSP). IR model infers on one Tesla V100 GPU. For HotpotQA, we use Wikipedia 2017 as the corpus, which is provided by [34] in full-wiki setting. For the other datasets, we use the large-scale passage collection built on Wikipedia as the corpus [13]. Baselines with information retrieval are in the same setting as SearChain. We reproduce all baselines on *gpt-3.5-turbo* following the settings in their papers. The maximum number of interaction rounds  $r_{max}$  is 5. The thresholds  $\alpha$  and  $\theta$  are

<sup>3</sup>https://openai.com/api/

<sup>&</sup>lt;sup>1</sup>Retrieve documents and provide them to LLM in a prompt.

<sup>&</sup>lt;sup>2</sup>Perform ToolFormer on gpt-3.5-turbo via in-context learning.

		Muti-F	Iop QA		Slot I	Filling	FC	LFQA
	HoPo	MQ	WQA	SQA	zsRE	T-REx	FEV.	ELI5
		Witho	out Informat	ion Retriev	al			
Direct Prompting	31.95	5.91	25.82	66.25	22.75	43.85	73.45	21.90
Auto-CoT	33.53	10.55	29.15	65.40	21.30	43.98	76.61	21.55
CoT	35.04	9.46	30.41	65.83	22.36	44.51	76.98	21.79
CoT-SC	36.85	10.02	32.68	70.84	24.74	46.06	77.15	22.05
Recite-and-answer	36.49	10.97	32.53	70.47	24.98	46.14	77.35	22.10
Self-Ask w/o IR	33.95	11.10	35.65	65.45	20.16	44.71	75.31	21.73
Least-to-Most	34.05	11.45	32.88	65.78	21.86	44.98	75.98	21.95
Plan-and-Solve	36.33	12.95	35.68	73.21	25.15	47.58	77.08	22.23
SearChain w/o IR	38.36	13.61	40.49	75.62	30.14	52.69	77.06	22.54
		Interactio	n with Info	mation Ret	rieval			
Direct Retrieval	34.09	10.22	30.01	66.78	52.29	59.28	78.25	23.40
ToolFormer	36.75	12.98	35.49	67.02	51.35	59.17	80.79	23.05
Self-Ask	40.05	14.28	39.58	67.65	50.51	59.12	79.41	23.25
Plan-and-Solve w/ IR	41.65	15.07	42.05	74.58	52.15	60.03	81.04	24.56
React $\rightarrow$ CoT-SC	43.15	15.49	40.36	70.43	53.27	60.42	80.59	24.05
Verify-and-Edit	44.03	15.57	40.83	71.09	53.95	61.10	80.67	23.80
Tree-of-Thought w/ IR	50.65	15.61	42.49	72.55	54.88	62.40	81.03	24.20
DSP	51.97	15.83	43.52	72.41	54.35	61.32	80.65	23.46
SearChain	56.91	17.07	46.27	76.95	57.29	65.07	81.15	25.57
- w/o Verification	46.11	14.70	42.67	75.98	43.58	55.46	78.79	22.98
- w/o Completion	53.05	15.86	43 64	76 53	45 78	56.03	80.03	25.02

Table 1: Performance of SearChain and baselines on complex knowledge-intensive tasks. Bold denotes the best result in different settings. FC: Fact Checking, LFQA: Long-Form QA. Metric for LFQA: ROUGE-L. Metric for others: cover-EM.

set as 0.35 and 1.5 respectively. As for the selection of confidence threshold ( $\theta$ ), we initialize the initial value of the confidence threshold (1.0) based on prior knowledge and gradually increase the value with a step size of 0.1. We validate the F1-score (a comprehensive metric of the Recall and Precision of judging whether the passage can answer the question) on the mixed open-domain OA datasets (NQ, TriviaQA, WebQ, and TREC) after each value change. We find that when the confidence threshold is 1.5, the highest F1-score can be achieved so we set the confidence threshold as 1.5. As for the selection of ROUGE threshold ( $\alpha$ ), we determine this value by observing the ROUGE relationship between the generated text and the ground truth in the few examples in in-context learning. Our further experiments show that when the value range of ROUGE threshold is between 0.3 and 0.5, the performance change on ELI5 is not obvious. Details of prompts and experiments are introduced in Section A.4 of Appendix.

# 4.2 Main Results

Performance of SearChain and baselines on complex knowledgeintensive tasks are shown in Table 1.

(1) Effect of Chain-of-Query. CoQ is the reasoning chain for
complex questions in SearChain. We compare it with recent competitive baselines in the setting without IR. SearChain w/o IR outperforms all baselines based on CoT (CoT, Auto-CoT, CoT-SC and
Recite-and-answer), which indicates that focusing on constructing a global reasoning chain consisting of sub-questions is better
than just giving intermediate reasoning results. SearChain w/o IR

outperforms Self-Ask w/o IR and Least-to-Most, which indicates that it is more effective to focus on constructing a global reasoning chain at each inference (global perspective) than generating and answering sub-questions step by step (local perspective).

(2) Effect of interaction with IR. In the setting with interaction with IR, SearChain again outperforms all the baselines. The paradigm of first generating global CoQ, and then IR interacting with each node of CoQ ensures the coherence of LLM reasoning. This solves the problem in Self-Ask, DSP and React. Besides, SearChain decouples the knowledge of LLM and IR. IR judges whether to provide information to LLM according to the confidence and the flag of the node on CoQ, which effectively alleviates misleading LLM. Last but not least, baselines reason in the one-dimensional chain. They cannot dynamically modify the reasoning path from a chain to node-identify Depth-first Search on a tree, which enables LLM dynamically modify the reasoning direction.

# 4.3 Analysis

In this section, we discuss and compare the advantages of SearChain compared to baseline in detail. First, we analyze the source of the knowledge of SearChain in solving complex questions. Second, while we analyze the positive effect of IR on LLM in solving difficult questions, we also demonstrate that our method can better mitigate the negative effect of IR on LLM. Third, we show the advantages of SearChain compared to baselines in terms of reasoning and tracing capabilities. Last but not least, we perform efficiency analysis to

Anon.

Table 2: Distribution of knowledge sources.

Knowledge Src.	HoPo	MQ	WQA	SQA
LLM	74.56%	78.83%	75.83%	94.98%
Corrected by IR	20.94%	14.60%	18.96%	2.78%
Completed by IR	4.50%	6.57%	5.21%	2.24%

### Table 3: Positive and negative effects of IR on LLM.

(a)	Accuracy	on $\mathbb{S}_{IR}$	and $\mathbb{S}$	(positive	effect	(1).
-----	----------	----------------------	------------------	-----------	--------	------

	HoPo	MÇ	2	WQA	SQA
w/o IR (S)	38.36	13.	61	40.49	75.62
w/o IR ( $\mathbb{S}_{IR}$ )	31.38	10.	20	32.60	68.96
w IR ( $\mathbb{S}_{IR}$ )	60.86	18.	49	50.52	78.42
(b) Percentage that IR misleads LLM (negative effect $\downarrow$ ).					
		HoPo	MQ	WQA	SQA
Self-Ask		15.76	14.32	25.76	10.29
React		17.68	15.22	25.99	10.03
Plan-and-Solve w/ IR		16.42	15.25	22.31	7.59
Verify-and-Edit		9.78	10.75	16.44	6.52
Tree-of-Thought w/ IR		12.07	13.25	20.52	8.46
DSP		14.72	14.03	24.31	9.22
SearChain		6.33	6.50	12.71	5.31

show our method significantly improves task performance with no significant increase in time consumption.

4.3.1 **Knowledge Decoupling**. We analyze the knowledge sources on the four multi-hop QA datasets. Specifically, we classify knowledge sources into three categories: (1) knowledge of LLM, (2) knowledge that corrected by IR in verification, and (3) knowledge that LLM does not know and is provided by IR in completion. We use node of ToR as the statistical granularity to calculate the percentage of nodes from these three sources in the total nodes respectively. The experimental results are shown in Table 2. It is worth noting that even though most of the knowledge comes from LLM, this knowledge is also verified by IR. IR only corrects the inconsistent answer given by LLM when it is confident enough and provides LLM with the missing knowledge, which alleviates the negative effect of IR on LLM and improves utilization of retrieved information. On StrategyQA, LLM has memorized most knowledge that IR can retrieve, so IR provides less knowledge than other datasets.

7454.3.2 Positive and Negative Effects of IR on LLM. (1) Positive.746In SearChain, IR can identify the trouble of LLM and effectively747help LLM to correct the answers and acquire missing knowledge.748We select the questions  $(\mathbb{S}_{IR})$  that IR helps to correct or provide749knowledge from the datasets used in Table 1 (\$\$) and evaluate the750accuracy of SearChain on  $\mathbb{S}_{IR}$ . We also evaluate the accuracy of751SearChain w/o IR on  $\mathbb{S}_{IR}$ . The results in Table 3(a) show that w/o752IR performs worse on  $\mathbb{S}_{IR}$  than on  $\mathbb{S}$ , which indicates that LLM753does have trouble with the questions that require IR help. w/ IR

 Table 4: Number of reasoning steps. SearChain tries more for unsolvable sub-questions to achieve better accuracy.

	2-hop	3-hop	4-hop	Accuracy
CoT	2.25	2.23	2.19	35.04
Self-Ask w/o IR	2.04	2.21	2.15	33.95
Least-to-Most	2.52	2.68	2.70	34.05
SearChain w/o IR	4.16	4.66	5.06	38.36

	Who founded what Robert J Chassell is a member of?
-	► CoT
	It is not clear what entity Robert J Chassell is a member of. The question cannot be answered. $ ightarrow$
•	Self-Ask w/o Retrieval
	Follow up: Who is Robert J Chassell? Intermediate answer: Robert J Chassell is a programmer and author. Follow up: What organization is Robert J Chassell a member of? Intermediate answer: The organization that Robert J Chassell is a member of is not specified. X
•	SearChain w/o Retrieval
	[Query 1]: What is Robert J Chassell a member of? [Answer 1]: There is no information about what organization or group Robert J Chassell is a member of. [Query 2]: Can we find any information about Robert J Chassell's affiliations or associations? [Answer 2]: Yes, Robert J Chassell is identified as a member of the Free Software Foundation. [Query 3]: Who founded the Free Software Foundation? [Answer 3]: The Free Software Foundation was founded by Richard Stallman. [Final Content]: Richard Stallman founded the Free Software Foundation [3], which Robert J Chassell is a member of [2]. So the final answer is Richard Stallman. $\checkmark$

# Figure 4: Case study of the difference between SearChain and baselines for unsolvable sub-questions.

performs better on  $S_{IR}$ , which indicates that IR effectively identifies and solves the trouble of LLM.

(2) Negative. We point out the risk of IR misleading LLM when there is a conflict in the knowledge of IR and LLM (Section 1). We select the questions ( $\mathbb{S}_t$ ) that LLM can give the correct answers to and count the percentage that LLM gives incorrect answers after adding IR on  $\mathbb{S}_t$ . Table 3(b) shows SearChain effectively mitigates the negative effect of IR on LLM. It is because that SearChain uses the confidence of IR and the information of CoQ to judge whether to correct LLM or provide LLM with its missing knowledge.

*4.3.3* **CoQ vs Baselines in Reasoning.** CoQ performs better on reasoning complex questions than the baselines. In addition to Table 1, we further analyze the reasoning ability from two aspects:

(1) Number of Reasoning Steps. We analyze the number of reasoning steps in different methods in the setting without IR. We conduct the experiment on Musique because Musique has more complex questions. Table 4 shows the average number of reasoning steps on questions with different hops. Our method has more reasoning steps, and the number of reasoning steps increases with the hops of the question. This shows that our method has a better perception of the complexity of the questions.

(2) Solving Difficult Sub-questions. The baselines focus on solving local sub-questions while ignoring the global planning of the reasoning chain. This leads LLM to tend to stop reasoning rather than try more when a sub-question cannot be solved. In our method, LLM acts as a commander that plans a global reasoning chain that can solve the complex question, when a sub-question cannot be solved, even without the help of IR, LLM can try to further decompose or rewrite the sub-question to continue reasoning. It is because our method focuses on building a global chain that can solve the

	1					
What team did the recipient of	What kind of dinosaur	does Raymond Ochoa	In what London botanical garden can you find the			
the 2007 Brownlow Medal play for?	voice in the Pixar film	voice in the Pixar film "The Good Dinosaur"?		fungus described by Elise Maud Wakefield as "Lepiota"?		
🖉 SearChain 🛛 🖒 New Bing	🖇 SearChain	New Bing	🖉 SearChain		New Bing	
Jimmy Bartel was the reci- The recipient of the 2007	Raymond Ochoa is an	Raymond Ochoa vo-	Elise Maud Wakefield	was an English	I'm sorry, but I couldn't find any	
pient of the 2007 Brown-low Brownlow Medal was	American child actor	ices the character of	mycologist and plant	pathologist [1].	information about a fungus described by	
Medal [1] and played Jimmy Bartel who	and voice actor [1]. He	Arlo, an Apatosaurus	"Lepiota" is a genus of	fungi [2]. Kew	Elsie Maud Wakefield as "Lepiota" in a	
Australian rules football played for the Geelong	has played various roles	in the Pixar film "The	Gardens is known for its fur	ngi collection [3],	London botanical garden. I did find that	
professionally [2]. He played Football Club. He won	Arlo in "The Good	Good Dinosaur" [1]	such as "Lepiota viridigleba	" [4]. So the final	Elsie Maud Wakefield was a mycologist at	
for the Geelong Football the medal by polling	Dinosaur" [2]. Arlo is		answer is Kew Gardens.		the Royal Botanic Gardens, Kew from	
Club in the Australian twenty-nine votes during	an Apatosaurus [3]. So	[1]:(Amazon.com: The			1910 to 1951 and was a specialist in	
Football League [2]. So the the 2007 AFL season [1].	the final answer is	Good Dinosaur : Ray-	[1]: Elsie Maud Wakefie	ld, was an English	Basidiomycetes and an international	
final answer is Geelong Is there anything else	Apatosaurus.	mond Ochoa)	mycologist and plant pathologist	ogist	authority on Aphyllophorales [1,2].	
Football Club. you would like to know	[1], Dermand Oshea is		[2]: Lepiota viridigleba i	s a species of seq-		
about this topic?	an American child actor		uestrate fungus in the family	Agaricaceae	[1]: Elsie Maud Wakefield, was an	
[1]: Jimmy Bartel won the	and voice actor		[3] Kew Gardens is a b	otanical garden in	English mycologist and plant	
medal during the 2007 AFL [1]: Jimmy Bartel	[2]:Raymond Ochoa		southwest London that hous	es the "largest and	pathologist	
season won the medal by polling	played, Arlo, in the		most diverse botanical	and mycological	[2]: Elsie Wakefield was a specialist in	
[2]: Jimmy Bartel is a twenty-nine votes during	Pixar film "The Good		collections in the world"		Basidiomycetes and an international auth-	
former Australian rules the 2007 AFL season	Dinosaur"		[4] The collection of fung	gi at Kew Gardens	ority on Aphyllophorales	
footballer plays for the	saurus named Arlo		includes species such as	"Lepiota viri-		
Geelong Football Club	(Ochoa)		digleba"			

Figure 5: Case study of SearChain and New Bing in marking references to supporting documents.

Table 5: Evaluation of traceability.

	SKC	AMP
New Bing	1.143	0.45
SearChain	2.882	0.80

complex question (global perspective), rather than answering or generating the sub-questions step by step (local perspective). Case study in Figure 4 shows that CoT and Self-Ask stop the reasoning while SearChain continues reasoning by rewriting the sub-question. More reasoning steps in Table 4 also support that SearChain can try more for difficult sub-questions. More case studies are shown in Section A.1.2 of Appendix.

4.3.4 SearChain vs New Bing in Tracing. We compare the performance of SearChain and New Bing<sup>4</sup> in marking references for generated contents via case study (Figure 5). We further propose two metrics to evaluate the Scope of Knowledge Coverage and Accuracy of Marking Position to show traceability more intuitively: • Scope of Knowledge Coverage (SKC) [0, +]: The number of knowledge items marked with supporting documents in the gener-ated contents. (statistics)

Accuracy of Marking Position (AMP) [0, 1]: The accuracy of
 the position of the reference marks. That is, whether the references
 are correctly marked on the sub-fragments for the corresponding
 knowledge in the generated contents. (human evaluation)

We introduce three humans with master's degrees to participate
in our human evaluation and the results are shown in Table 5.
SearChain can mark references for each knowledge involved in
the reasoning process (i.e., correct nodes of CoQ) in a fine-grained
manner. While the references given by New Bing do not cover all
of the knowledge and cannot be marked on the correct position.
More case studies are shown in Section A.1.1 of Appendix.

4.3.5 *Efficiency Analysis.* We analyze the running efficiency between SearChain and baselines on the number of words in the

<sup>4</sup>https://www.bing.com/new

## 

Table 6: Efficiency analysis.

	$\#n\downarrow$	$\#m\downarrow$	$\#r\downarrow$	$t(s) \downarrow$	Perf. (Avg) $\uparrow$
Self-Ask	401	63	2.19	6.63	46.73
Plan-and-Solve w/ IR	450	71	1	6.05	48.89
React $\rightarrow$ CoT-SC	938	110	2.35	8.25	48.47
Verify-and-Edit	565	307	2.40	13.90	48.88
Tree-of-Thought w/ IR	622	341	2.29	13.28	50.47
DSP	1759	155	2.15	10.47	50.44
SearChain	390	189	2.21	8.52	53.29

input (*n*) and output (*m*) text of LLM, number of rounds of interaction between LLM and IR (*r*) and overall running time (*t*). Table 6 shows our method significantly improves task performance with no significant increase in time consumption. Most baselines also require multiple rounds of interaction between IR and LLM.

# 5 CONCLUSION

In this paper, we pointed out the challenges of introducing IR into LLM from the perspectives of reasoning and knowledge. We then proposed a novel framework named SearChain to enable IR and LLM to interact with each other effectively. SearChain not only stimulates the knowledge-reasoning ability of LLM but also uses IR to provide the knowledge that LLM really needs based on the external knowledge base, which improves both accuracy and credibility. Besides, SearChain can mark references to supporting documents for the knowledge involved in the generated contents, which improves the traceability of the contents. In addition, the interaction between IR and LLM in SearChain transforms the reasoning path from a chain to node-identify Depth-first Search on a tree, which enables LLM to dynamically modify the reasoning direction. Experimental results on complex knowledge-intensive tasks show that SearChain performs better than all baselines.

Anon

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

1022

1023

1024

1025

1026

1027

1028

1029

1030

1031

1032

1033

1034

1035

1036

1037

1038

1039

1040

1041

## 929 **REFERENCES**

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

986

- Razvan Azamfirei, Sapna R Kudchadkar, and James Fackler. 2023. Large language models and the perils of their hallucinations. *Critical Care* 27, 1 (2023), 1–2.
- [2] Yejin Bang, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wile, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, Quyet V. Do, Yan Xu, and Pascale Fung. 2023. A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. *CoRR* abs/2302.04023 (2023). arXiv:2302.04023 https://doi.org/10.48550/arXiv.2302.04023
- [3] Sebastian Borgeaud, Arthur Mensch, Jordan Hoffmann, et al. 2022. Improving Language Models by Retrieving from Trillions of Tokens. In International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA (Proceedings of Machine Learning Research, Vol. 162). PMLR, 2206–2240. https://proceedings.mlr.press/v162/borgeaud22a.html
  - [4] Andrew Drozdov, Nathanael Schärli, Ekin Akyürek, Nathan Scales, Xinying Song, Xinyun Chen, Olivier Bousquet, and Denny Zhou. 2023. Compositional Semantic Parsing with Large Language Models. In *The Eleventh International Conference* on Learning Representations. https://openreview.net/forum?id=gJW8hSGBys8
  - [5] Hady Elsahar, Pavlos Vougiouklis, Arslen Remaci, et al. 2018. T-REx: A Large Scale Alignment of Natural Language with Knowledge Base Triples. In Proceedings of the 2018 Conference on LREC. European Language Resources Association (ELRA), Miyazaki, Japan. https://aclanthology.org/L18-1544
  - [6] Angela Fan, Yacine Jernite, Ethan Perez, et al. 2019. ELI5: Long Form Question Answering. In Proceedings of the 2019 Conference on ACL. Association for Computational Linguistics, Florence, Italy, 3558–3567. https://doi.org/10.18653/v1/P19-1346
- [7] Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark, and Tushar Khot. 2023. Complexity-Based Prompting for Multi-step Reasoning. In *The Eleventh International Conference on Learning Representations*. https://openreview.net/forum? id=yf1icZHC-19
- [8] Mor Geva, Daniel Khashabi, Elad Segal, Tushar Khot, Dan Roth, and Jonathan Berant. 2021. Did Aristotle Use a Laptop? A Question Answering Benchmark with Implicit Reasoning Strategies. *Transactions of the Association for Computational Linguistics (TACL)* (2021).
- [9] Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat, and Ming-Wei Chang. 2020. REALM: Retrieval-Augmented Language Model Pre-Training. CoRR abs/2002.08909 (2020). arXiv:2002.08909 https://arxiv.org/abs/2002.08909
- [10] Xanh Ho, Anh-Khoa Duong Nguyen, Saku Sugawara, and Akiko Aizawa. 2020. Constructing A Multi-hop QA Dataset for Comprehensive Evaluation of Reasoning Steps. In Proceedings of the 2020 Conference on COLING. International Committee on Computational Linguistics, Barcelona, Spain (Online), 6609–6625. https://doi.org/10.18653/v1/2020.coling-main.580
- [11] Gautier Izacard and Edouard Grave. 2020. Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering. *CoRR* abs/2007.01282 (2020). arXiv:2007.01282 https://arxiv.org/abs/2007.01282
- [12] Nikhil Kandpal, Haikang Deng, Adam Roberts, Eric Wallace, and Colin Raffel. 2022. Large Language Models Struggle to Learn Long-Tail Knowledge. arXiv:2211.08411 [cs.CL]
- [13] Vladimir Karpukhin, Barlas Oguz, Sewon Min, et al. 2020. Dense Passage Retrieval for Open-Domain Question Answering. In *Proceedings of the 2020 Conference on EMNLP*. Association for Computational Linguistics, Online, 6769–6781. https: //doi.org/10.18653/v1/2020.emnlp-main.550
- [14] Omar Khattab, Keshav Santhanam, Xiang Lisa Li, et al. 2023. Demonstrate-Search-Predict: Composing retrieval and language models for knowledge-intensive NLP. arXiv:2212.14024 [cs.CL]
- [15] Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large Language Models are Zero-Shot Reasoners. In Advances in Neural Information Processing Systems, Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (Eds.). https://openreview.net/forum? id=e2TBb5y0yFf
- [16] Omer Levy, Minjoon Seo, Eunsol Choi, and Luke Zettlemoyer. 2017. Zero-Shot Relation Extraction via Reading Comprehension. In *Proceedings of the 2017 Conference on CoNLL*). Association for Computational Linguistics, Vancouver, Canada, 333–342. https://doi.org/10.18653/v1/K17-1034
- [17] Patrick S. H. Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In Proceedings of the 2020 Conference on NeurIPS.
- [18] Chin-Yew Lin. 2004. ROUGE: A Package for Automatic Evaluation of Summaries. In Text Summarization Branches Out. Association for Computational Linguistics, Barcelona, Spain, 74–81. https://aclanthology.org/W04-1013
- [19] Sewon Min, Weijia Shi, Mike Lewis, Xilun Chen, Wen-tau Yih, Hannaneh Hajishirzi, and Luke Zettlemoyer. 2022. Nonparametric Masked Language Modeling. *CoRR* abs/2212.01349 (2022). https://doi.org/10.48550/arXiv.2212.01349 arXiv:2212.01349
- [20] Fabio Petroni, Aleksandra Piktus, Angela Fan, et al. 2021. KILT: a Benchmark for Knowledge Intensive Language Tasks. In *Proceedings of the 2021 Conference on*

NAACL. Association for Computational Linguistics, Online, 2523–2544. https://doi.org/10.18653/v1/2021.naacl-main.200

- [21] Ofir Press, Muru Zhang, Sewon Min, Ludwig Schmidt, Noah A. Smith, and Mike Lewis. 2023. Measuring and Narrowing the Compositionality Gap in Language Models. https://openreview.net/forum?id=PUwbwZJz9dO
- [22] Hongjing Qian, Yutao Zhu, Zhicheng Dou, et al. 2023. WebBrain: Learning to Generate Factually Correct Articles for Queries by Grounding on Large Web Corpus. arXiv:2304.04358 [cs.CL]
- [23] Corby Rosset, Chenyan Xiong, Minh Phan, et al. 2020. Knowledge-Aware Language Model Pretraining. *CoRR* abs/2007.00655 (2020). arXiv:2007.00655 https://arxiv.org/abs/2007.00655
- [24] Keshav Santhanam, Omar Khattab, Jon Saad-Falcon, Christopher Potts, and Matei Zaharia. 2022. ColBERTv2: Effective and Efficient Retrieval via Lightweight Late Interaction. In Proceedings of the 2022 Conference on NAACL. Association for Computational Linguistics, Seattle, United States, 3715–3734. https://doi.org/10. 18653/v1/2022.naacl-main.272
- [25] Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. Toolformer: Language Models Can Teach Themselves to Use Tools. arXiv:2302.04761 [cs.CL]
- [26] Aarohi Srivastava, Abhinav Rastogi, Abhishek Rao, Abu Awal Md Shoeb, Abubakar Abid, Adam Fisch, Adam R. Brown, Adam Santoro, Aditya Gupta, Adrià Garriga-Alonso, et al. 2022. Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models. arXiv:2206.04615 [cs.CL]
- [27] Zhiqing Sun, Xuezhi Wang, Yi Tay, Yiming Yang, and Denny Zhou. 2023. Recitation-Augmented Language Models. In *The Eleventh International Confer*ence on Learning Representations. https://openreview.net/forum?id=-cqvvvb-NkI
- [28] Robert Tarjan. 1971. Depth-first search and linear graph algorithms. In 12th Annual Symposium on Switching and Automata Theory (swat 1971). 114–121. https://doi.org/10.1109/SWAT.1971.10
- [29] James Thorne, Andreas Vlachos, Christos Christodoulopoulos, and Arpit Mittal. 2018. FEVER: a Large-scale Dataset for Fact Extraction and VERification. In Proceedings of the 2018 Conference on NAACL. Association for Computational Linguistics, New Orleans, Louisiana, 809–819. https://aclanthology.org/N18-1074
- [30] Harsh Trivedi, Niranjan Balasubramanian, Tushar Khot, and Ashish Sabharwal. 2022. MuSiQue: Multihop Questions via Single-hop Question Composition. *Transactions of the Association for Computational Linguistics* 10 (2022), 539–554. https://doi.org/10.1162/tacl\_a\_00475
- [31] Lei Wang, Wanyu Xu, Yihuai Lan, et al. 2023. Plan-and-Solve Prompting: Improving Zero-Shot Chain-of-Thought Reasoning by Large Language Models. arXiv:2305.04091 [cs.CL]
- [32] Xuezhi Wang, Jason Wei, Dale Schuurmans, et al. 2023. Self-Consistency Improves Chain of Thought Reasoning in Language Models. In *The Eleventh International Conference on Learning Representations*. https://openreview.net/forum? id=1PL1NIMMrw
- [33] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed H. Chi, Quoc Le, and Denny Zhou. 2022. Chain of Thought Prompting Elicits Reasoning in Large Language Models. *CoRR* abs/2201.11903 (2022). arXiv:2201.11903 https://arxiv.org/abs/2201.11903
- [34] Zhilin Yang, Peng Qi, Saizheng Zhang, et al. 2018. HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering. In *Proceedings of the 2018 Conference on EMNLP*. Association for Computational Linguistics, Brussels, Belgium, 2369–2380. https://doi.org/10.18653/v1/D18-1259
- [35] Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of Thoughts: Deliberate Problem Solving with Large Language Models. arXiv:2305.10601 [cs.CL]
- [36] Shunyu Yao, Jeffrey Zhao, Dian Yu, et al. 2023. ReAct: Synergizing Reasoning and Acting in Language Models. In *The Eleventh International Conference on Learning Representations*. https://openreview.net/forum?id=WE\_vluYUL-X
- [37] Da Yin, Li Dong, Hao Cheng, Xiaodong Liu, Kai-Wei Chang, Furu Wei, and Jianfeng Gao. 2022. A Survey of Knowledge-Intensive NLP with Pre-Trained Language Models. arXiv:2202.08772 [cs.CL]
- [38] Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah Goodman. 2022. STaR: Bootstrapping Reasoning With Reasoning. In Advances in Neural Information Processing Systems, Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (Eds.). https://openreview.net/forum?id=\_3ELRdg2sgI
- [39] Zhuosheng Zhang, Aston Zhang, Mu Li, and Alex Smola. 2023. Automatic Chain of Thought Prompting in Large Language Models. In The Eleventh International Conference on Learning Representations. https://openreview.net/forum? id=5NTt8GFjUHkr
- [40] Ruochen Zhao, Xingxuan Li, Shafiq Joty, Chengwei Qin, and Lidong Bing. 2023. Verify-and-Edit: A Knowledge-Enhanced Chain-of-Thought Framework. In Proceedings of the 2023 Conference on ACL. Association for Computational Linguistics, Toronto, Canada, 5823–5840. https://doi.org/10.18653/v1/2023.acl-long.320
- [41] Wayne Xin Zhao, Kun Zhou, Junyi Li, et al. 2023. A Survey of Large Language Models. arXiv:2303.18223 [cs.CL]
- [42] Denny Zhou, Nathanael Schärli, Le Hou, et al. 2022. Least-to-Most Prompting Enables Complex Reasoning in Large Language Models. arXiv:2205.10625 [cs.AI]

#### A APPENDIX

# A.1 Case Study

In this section, we compare the performance of SearChain and New Bing <sup>5</sup> in adding references to supporting documents for generated content via case study. We also use case study to further analyze why CoQ has stronger reasoning ability than Baselines. 

A.1.1 Case Study for SearChain vs New Bing in Tracing. We com-pare the performance of SearChain and New Bing in marking references for generated contents via case study (Figure 6  $\sim$  7). SearChain can mark references for each knowledge involved in the reasoning process (i.e., each correct node of CoQ) in a more fine-grained manner. While references given by New Bing do not cover all of the knowledge, and in some cases New Bing cannot find the knowledge. SearChain provides a novel perspective that decomposes complex multi-step knowledge-intensive tasks into multiple single-step knowledge reasoning problems, retrieving the supporting documents of knowledge for each step of reasoning, and organizing these reasoning steps with their reference marks as final generated content. This enables the supporting documents to cover every knowledge involved in the generated content, which enhances the traceability of the generated content. 

10/9		
1008	The movies The Boatniks and The Great Locomotive	Chase were both made by which production company?
1069	🖇 SearChain	b New Bing
1070	The production company for The Boatniks movie is Walt	Both "The Boatniks" and "The Great Locomotive Chase"
1070	Disney Productions [1]. The production company for	were made by Walt Disney Productions [1].
1071	The Great Locomotive Chase movie is Walt Disney	
1070	Productions [2]. So answer is Walt Disney Productions.	[1]: The Boatniks is a 1970 American comedy film
1072		directed by Norman Tokar and starring Robert Morse,
1073	[1]: The Boatniks is a 1970 American comedy film	Stefanie Powers, Don Ameche and Phil Silvers. It was
	starring Robert Morse, Stefanie Powers, Don Ameche	made by wait Disney Productions, released by Buena
1074	and Phil Silvers made by wait Disney Productions,	Vista Distribution
1075	[2] The Great Eccomotive Chase is a 1950 wait	
1074	What is the 2010 population of the city 2.1 m	iles southwest of Marietta Air Force Station?
1076	SoarChain	New Bing
1077	∠ Searchain The gity 2.1 miles couthwast of Mariatta Air Force.	I'm sorry but I couldn't find any information about a
1080	Station is Smyrna [1] The population of Smyrna in 2010	Marietta Air Force Station or a city located 2.1 miles
1078	was 51, 271 [2]. So the final answer is 51,271	southwest of it. There is a Dobbins Air Reserve Base
1079		located southeast of Marietta, Georgia, but I couldn't
	[1]: Marietta Air Force Station It is located 2.1	find any information about the population of a city
1080	miles northeast of Smyrna, Georgia	located 2.1 miles southwest of it [1].
1081	[2]: Smyrna is a city northwest of the neighborhoods	
	of Atlanta. As of the 2010 census, the city had a	<ol> <li>Dobbins Air Reserve Base is a United States Air</li> </ol>
1082	population of 51,271	Force reserve air base located in Marietta, Georgia, a
1083		suburb about 20 miles (32 km) northwest of Atlanta
	Which lead actor/actress in War Chhod Na Yaar has	s also acted in Bengali and English-language films?
1084	🖉 SearChain	New Bing
1085	The lead actors/actresses in War Chhod Na Yaar are	I'm sorry, but I couldn't find any information about any
	Sharman Joshi, Soha Ali Khan, and Javed Jaffrey [1].	lead actor or actress in War Chhod Na Yaar who has also
1086	Soha Ali Khan acted in Bengali and English-language	acted in Bengali and English-language films. The lead
1087	tilms according [2]. So the answer is Sona Ali Khan.	actors in War Chhod Na Yaar are Sharman Joshi, Soha
	[1]. Jacki Saka Ali Khan Jawad Jaffam, Samiai	Ali Khan, and Javed Jaffrey [1]. Is there anything else
1088	Mishra Dalin Tabil and Mukul Dev who play important	you would like to know about this topic?
1089	characters in the film	[1] War Chhod Na Yaar lead actors. Sharman Joshi
	[2]: Soha Ali Khan Khemu (born Soha Ali Khan , 4	Soha Ali Khan, Javed Jaffery, Sanjai Mishra. Dalip Tahil.
1090	October 1978) is an Indian film actress, she has also	and Mukul Dev.
1091	acted in Bengali and English-language films.	
1000	How large is the shopping mall where	e KGOT radio station has its studios ?
1092	🖇 SearChain	New Bing
1093	The studios of KGOT radio station are located in the	I'm sorry, but I couldn't find any information about the
	Dimond Center [1], which is a 728,000 ft <sup>2</sup> (67,000 m <sup>2</sup> )	size of the shopping mall where KGOT radio station has
1094	shopping mall in Anchorage, Alaska [2]. The size of the	its studios. KGOT is a commercial FM radio station in
1095	shopping mall where KGOT radio station has its studios	Anchorage, Alaska, and it broadcasts from studios on
	is 728,000 ft² (67,000 m²). So the answer is 728,000 ft² .	East Dimond Boulevard in the Dimond Center [1].
1096		
1097	[1]:KGOT (101.3 FM) is a commercial Top 40 (CHR)	[1]: KGOT (101.3 MHz) is a commercial FM radio
1071	radio station Dimond Center.	station on East Dimond Boulevard in the Dimond
1098	[2]:Dimond Center is a 728,000 ft <sup>2</sup> (67,000 m <sup>2</sup> )	Center.
	snopping man in Anchorage, Alaska	
1099		

Who is the author of the 1993 prod	uction Madge Ryan participated in?
8 SearChain	New Bing
Madge Kyan was known for her role in Medea in the year 1993 [1]. Medea was written by Euripides [2]. So	1 m sorry but I couldn't find any information about the author of the 1993 production Madge Ryan participated
the final answer is Euripides.	in. Could you please provide more information or
-	context so that I can help you better?
[1] Madge Ryan Madge Winifred Ryan was an	
Australian actress, known for her stage roles in London productions of "Entertaining Mr Sloape" (1964)	
"Philadelphia, Here I Come" (1967), and "Medea"	
(1993)	
[2] Medea (Seneca)   Medea (Seneca) Medea is a	
"tabula crepidata" (Roman tragedy with Greek subject)	
Who is the psychologist who wrote about the disciple	of the person known as the "Wild Woman of Duttl-"
SearChain	New Bing
Mary MacLane is known as the "Wild Woman of Buttle"	I'm sorry, but I couldn't find any information about the
[1]. James G. Kiernan wrote about Mary MacLane's	disciple of the person known as the "Wild Woman of
disciple Viola Larsen [2]. So, the final answer is James G.	Buttle" or the psychologist who wrote about them. Could
Kieman.	you provide more details or perhaps rephrase your
[1] Mary MacLane was a controversial Canadian-	quesuon:
born. MacLane was known as the "Wild Woman of	
Butte"	
[2] James G. Kiernan was an American psychologist, Kiernan went on to write of a variety of topics, e.g. More	
MacLane's disciple Viola Larsen	
The Dogwoman telemovies were cre	ated by an actress born in what year?
🖇 SearChain	New Bing
Magda Szubanski played the role in Dogwoman	I'm sorry, but I couldn't find any information about the
April 12, 1961 [3]. So the answer is 1961.	Could you provide more details or perhaps rephrase your
	question?
[1] Morice had a starring role in the series of three	
"Dogwoman" telemovies with Magda Szubanski	
[2] magua Szubanski In 1999, Szubanski created, wrote, co-produced and starred as Margaret O'Halloran	
in the "Dogwoman" series of TV films	
[3] Magda Szubanski was born 12 April 1961	
Jack Alfred Evans played with a football clu	b team that is based in what city in Australia?
2 · · · ·	1
SearChain	New Bing
SearChain Jack Alfred Evans played with St Kilda football club team [1] which is located in Victoria [2] and	New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football Lagrance
≫ SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is	Dew Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne.
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.	Deve Bing Jack Alfred Evans was an Australian rules footballer who played with SK Kilda in the Victorian Football League (VFL) [1]. SK Kilda Football Club is based in Melbourne, Victoria, Australia.
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.	New Bing     Jack Alfred Evans was an Australian rules footballer who     played with St Kilda in the Victorian Football League     (VFL) [1]. St Kilda Football Club is based in Melbourne,     Victoria, Australia.     [1] Jack Alfred Evang (hor: 11 October 1020)
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL	New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia. <ol> <li>Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in</li> </ol>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Austalian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> </ul>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> </ul>
<ul> <li>SearChain</li> <li>Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan of the metropolitan with the second se</li></ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> </ul>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia	New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian. [1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia "The Best of LeAnn Rimes" contained a song used as ta for the metropolitan and the field of the field o	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> </ul>
<ul> <li>SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda, Victoria of Australia is located in Alma area of Melbourne, Victoria, Australia</li> <li>"The Best of LeAnn Rimes" contained a song used as t</li></ul>	New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia. [1] Jack Alfred Evans (born 11 October 1930) was an Australian rules football club op layed with St Kilda in the Victorian Football League (VFL). he theme song for a film. What is the name of the film? New Bing Can't Eight the Mosolicht <sup>to</sup> by LaAre Pience
<ul> <li>SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria, Australia </li> <li>The Best of LeAnn Rimes" contained a song used as t SearChain The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme </li> </ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Covete Ugly 11).</li> </ul>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda victoria of Australia is located in Alma Road [3] St Kilda Anne mess" contained a song used as t S SearChain The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is	New Bing     Jack Alfred Evans was an Australian rules footballer who     played with St Kilda in the Victorian Football League     (VFL) [1]. St Kilda Football Club is based in Melbourne,     Victoria, Australian     [1] Jack Alfred Evans (born 11 October 1930) was an     Australian rules footballer who played with St Kilda in     the Victorian Football League (VFL).     he theme song for a film. What is the name of the film?     New Bing     The song "Can't Fight the Moonlight" by LeAnn Rimes     was used as the theme song for the film Coyote Ugly [1].     Is there anything else you would like to know?
≥> SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia "The Best of LeAnn Rimes" contained a song used as t SearChain The album "The Best of LeAnn Rimes" includes the song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> </ul>
≥> SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia "The Best of LeAnn Rimes" contained a song used as t SearChain The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight</li> </ul>
<ul> <li>SearChain</li> <li>Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia</li> <li>"The Best of LeAnn Rimes" contained a song used as t SaraChain</li> <li>The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".</li> <li>[1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in Victoria Sara Chains and Sara Chan Rimes is a greatest hits album by American Coyote Ugly".</li> </ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing Was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
➢ searChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia "The Best of LeAnn Rimes" contained a song used as t SearChain The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly". [1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, "cleased in 2004, including "How Do I Live", "Can't Fight the	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL)[], St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly []. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
<ul> <li>SearChain</li> <li>Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda Nettoria of Australia is located in Alma Road</li> <li>[3] St Kilda Nettoria of Australia</li> <li>"The Best of LeAnn Rimes" contained a song used as t S SearChain</li> <li>The album "The Best of LeAnn Rimes" includes the song 'Cna't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".</li> <li>[1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004, including "How Do I Live", "Can't Fight the Moonlight" and ner duet with Ronan Keating, "Last</li> </ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria, Australia The Best of LeAnn Rimes" contained a song used as t SearChain The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly". [1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004,, including "How Do I Live", "Can't Fight the Moonlight" and her duct with Ronan Keating, "Last Thing on My Mind".	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
➢ searChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia "The Best of LeAnn Rimes" contained a song used as t SearChain The album "The Best of LeAnn Rimes" includes the song "Cart Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly". [1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004,, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind". [2] "Can't Fight the Moonlight" is a song writen by Diane Warren, recorded it was the theme song of the song	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
<ul> <li>SearChain</li> <li>Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda Nictoria of Australia is located in Alma Road</li> <li>[3] St Kilda Nictoria of Australia</li> <li>"The Best of LeAnn Rimes" contained a song used as t Song "Cant Fight the Moonlight" [1], which is the theme song of the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".</li> <li>[1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind".</li> <li>[2] "Can't Fight the Moonlight" is a song written by Diane Warren, recorded It was the theme song of the film "Coyote Ugly".</li> </ul>	New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian [1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL). he theme song for a film. What is the name of the film? New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know? [1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]
➢ SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda victoria of Australia "The Best of LeAnn Rimes" contained a song used as t	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>wylor Swift released on October 22, 2012?</li> <li>New Bing</li> </ul>
<ul> <li>≥ SearChain         Jack Alfred Evnas played with St Kilda football club         team [1], which is located in Victoria [2], and         specifically in Melbourne [3]. So the final answer is         Melbourne.         [1] Jack Evans was an Australian rules footballer who         played with St Kilda VFL         [2] St Kilda, Victoria of Australia is located in Alma         Road         [2] St Kilda, Victoria of Australia is located in Alma         Road         [2] St Kilda is an inner suburb of the metropolitan         area of Melbourne, Victoria, Australia         "The Best of LcAnn Rimes" contained a song used as t         SearChain         The album "The Best of LeAnn Rimes" includes the         song for the film "Coyote Ugly" [2]. So the answer is         "Coyote Ugly".         [1] The Best of LeAnn Rimes is a greatest hits album by         American country singer LeAnn Rimes, released in         2004,, including "How Do I Live", "Can't Fight the         Moonlight" and her duet with Ronan Keating, "Last         Thing on My Mind".         [2] "Can't Fight the Moonlight" is a song written by         Diane Warren, recorded It was the theme song of the         film "Coyote Ugly".         When did the tour begin for the album Taylor Swift released on October         </li> </ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda victoria of Australia is located in Alma Road [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia "The Best of LeAnn Rimes" contained a song used as t SearChain The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly". [1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004,, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind". [2] "Can't Fight the Moonlight" is a song written by Diane Warre, recorded It was the theme song of the film "Coyote Ugly". When did the tour begin for the album Ti SearChain The name of the album Taylor Swift released on October 22, 2012, is "Red" [1], "The Red Tour" was organized in	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>wift released on October 22, 2012?</li> <li>New Bing Taylor Swift released her fourth studio album, "Red," on October 22, 2012 [1]. However, I'm sorry but I couldn't</li> </ul>
≥ SearChain           Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.           [1] Jack Evans was an Australian rules footballer who played with St Kilda VEL           [2] St Kilda, Victoria of Australia is located in Alma Road           [3] St Kilda VEL           [2] St Kilda victoria of Australia is located in Alma Road           [3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia           The Best of LeAnn Rimes' contained a song used as t           SearChain           The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".           [1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004,, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind".           [2] "Can't Fight the Moonlight" is a song written by Diane Warren, recorded It was the theme song of the film "Coyote Ugly".           When did the tour begin for the album Ti SearChain           The name of the album Taylor Swift released on October 22, 2012, is "Red" [1]. "The Red Tour" was organized in support of the album Taylor Swift released on October 22, 2012, is "Red" [1]. "The Red Tour" was organized in Support of the album Taylor Swift released on October 22, 2012, is "Red" [1]. "The Red Tour" was organized in Suphopring hareaver is March 13. 2013. <td><ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing Taks on "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>wylor Swift released on October 22, 2012?</li> <li>New Bing Taylor Swift released her fourth studio album, "Red," on October 22, 2012 [1]. However, I'm sorry but I couldn't find the exact date when the tour for this album began. Is there any this rade are would like to know like to album segan.</li> </ul></td>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australia.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing Taks on "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>wylor Swift released on October 22, 2012?</li> <li>New Bing Taylor Swift released her fourth studio album, "Red," on October 22, 2012 [1]. However, I'm sorry but I couldn't find the exact date when the tour for this album began. Is there any this rade are would like to know like to album segan.</li> </ul>
<ul> <li>≥ SearChain</li> <li>Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda victoria of Australia is located in Alma Road</li> <li>[4] The Best of LeAnn Rimes" contained a song used as t</li></ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>waylor Swift released on October 22, 2012?</li> <li>New Bing Taylor Swift released her fourth studio album, "Red," on October 22, 2012 [1]. However, I'm sorry but I couldn't find the exact date when the tour for this album began. Is there anything else you would like to know?</li> </ul>
<ul> <li>≥&gt; SearChain</li> <li>≥&gt; SearChain</li> <li>&gt;&gt; SearChain</li> <li>&gt;&gt; So the final answer is Melbourne [3]. So the final answer is Melbourne.</li> <li>(1) Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>(2) St Kilda Victoria of Australia is located in Alma Road</li> <li>(3) St Kilda Victoria of Australia is located in Alma Road</li> <li>(3) St Kilda Victoria of Australia is located in Alma Road</li> <li>(2) St Kilda Victoria of Australia is located in Alma Road</li> <li>The Best of LeAnn Rimes" contained a song used as t So SearChain</li> <li>The album "The Best of LeAnn Rimes" includes the song 'Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".</li> <li>(1) The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004,, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind".</li> <li>(2) "Can't Fight the Moonlight" is a song written by Diane Warren, recorded It was the theme song of the film "Coyote Ugly".</li> <li>&gt;&gt; When did the tour begin for the album Ta So SearChain</li> <li>The name of the album Taylor Swift released on October 22, 2012, is "Red" [1], "The Red Tour" was organized in support of the album "Taylor Swift released on October 13, 2013 [3]. So, the final answer is March 13, 2013</li> <li>(1] Red (Taylor Swift album) released on October</li> </ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
<ul> <li>≥&gt; SearChain</li> <li>≥&gt; SearChain</li> <li>Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne.</li> <li>[1] Jack Evans was an Australian rules footballer who played with St Kilda VFL</li> <li>[2] St Kilda, Victoria of Australia is located in Alma Road</li> <li>[3] St Kilda victoria of Australia is located in Alma Road</li> <li>[3] St Kilda victoria of Australia is located in Alma Road</li> <li>[3] St Kilda is an inner suburb of the metropolitan area of Melbourne, Victoria, Australia</li> <li>"The Best of LcAnn Rimes" contained a song used as t S SearChain</li> <li>The album "The Best of LeAnn Rimes" includes the song "Can't Fight the Moonlight" [1], which is the theme song for the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly".</li> <li>[1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004,, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind".</li> <li>[2] "Can't Fight the Moonlight" is a song written by Diane Warre, recorded It was the theme song of the film "Coyote Ugly".</li> <li>When did the tour begin for the album Ta S SearChain</li> <li>When did the tour begin for the album Ta Joine Warre, the of the album Taylor Swift released on October 22, 2012, is "Red" [1], "The Red Tour" was organized in support of the album "Red" [2], and it started on March 13, 2013 [3]. So, the final answer is March 13, 2013.</li> <li>[1] Red (Taylor Swift album) released on October 22, 2012</li> </ul>	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> </ul>
	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian.</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimes was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>w Bing New Bing</li> <li>he w Bing Taylor Swift released her fourth studio album, "Red," on October 22, 2012 [1]. However, I'm sorry but I couldn't find the exact date when the tour for this album began. Is there anything else you would like to know?</li> <li>[1] Red (Taylor Swift album) released on October 22, 2012</li> </ul>
SearChain Jack Alfred Evans played with St Kilda football club team [1], which is located in Victoria [2], and specifically in Melbourne [3]. So the final answer is Melbourne. [1] Jack Evans was an Australian rules footballer who played with St Kilda VFL [2] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia is located in Alma Road [3] St Kilda, Victoria of Australia "The Best of LeAnn Rimes" contained a song used as t <i>§</i> SearChain The Best of LeAnn Rimes" includes the song "Cant Fight the Moonlight" [1], which is the theme song or the film "Coyote Ugly" [2]. So the answer is "Coyote Ugly". [1] The Best of LeAnn Rimes is a greatest hits album by American country singer LeAnn Rimes, released in 2004, including "How Do I Live", "Can't Fight the Moonlight" and her duet with Ronan Keating, "Last Thing on My Mind". [2] "Can't Fight the Moonlight" is a song written by Diane Warren, recorded It was the theme song of the film "Coyote Ugly". When did the tour begin for the album Tig SearChain The Red Tour "Mos organized in support of the album Taylor Swift released on October 22, 2012, is "Red" [1]. "The Red Tour" was organized in support of the album "Red" [2], and it started on March 13, 2013 [3]. So, the final answer is March 13, 2013. [1] Red (Taylor Swift album) released on October 22, 2012, in "Line Red Tour was the third concert tour by Taylor Swift. Launched in support of Swift's fourth studo album, "Red" "	<ul> <li>New Bing Jack Alfred Evans was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL) [1]. St Kilda Football Club is based in Melbourne, Victoria, Australian</li> <li>[1] Jack Alfred Evans (born 11 October 1930) was an Australian rules footballer who played with St Kilda in the Victorian Football League (VFL).</li> <li>he theme song for a film. What is the name of the film?</li> <li>New Bing The song "Can't Fight the Moonlight" by LeAnn Rimess was used as the theme song for the film Coyote Ugly [1]. Is there anything else you would like to know?</li> <li>[1] LeAnn Rimes - Can't Fight The Moonlight (Theme from "Coyote Ugly") [1080p]</li> <li>wey Bing Taylor Swift released on October 22, 2012?</li> <li>New Bing Taylor Swift released her fourth studio album, "Red," on October 22, 2012 [1]. However, I'm sorry but I couldn't find the exact date when the tour for this album began. Is there anything else you would like to know?</li> <li>[1] Red (Taylor Swift album) released on October 22, 2012</li> </ul>

# Figure 6: Case study for SearChain vs New Bing in Tracing.

<sup>5</sup>https://www.bing.com/new

The Duke Steps Out stars an actress who was ranked tenth on a list of grea	test female Hollywood stars in what yea
🖉 SearChain	🕨 New Bing
Joan Crawford starred in The Duke Steps Out [1], which was released on May 14, 1929 [2] The American Film Institute (AFI) ranked the greatest	I'm sorry, but I couldn't find an
female Hollywood stars in 1999 [3], and the top five were Katharine	TheDuke Steps Out who was ranke
Hepburn, Bette Davis, Audrey Hepburn, Ingrid Bergman, and Greta	tenth on a list of greatest fema
Garbo [4]. Joan Crawford was ranked tenth on the list [5]. So the answer	Hollywood stars [1]. Could you provid
is 1999.	more details or perhaps rephrase yo question?
[1] The Duke Steps Out The Duke Steps Out is directed by James	question
Cruze and starring William Haines and Joan Crawford	[1]: AFI's 100 Years100 Stars is the
[2] The Duke Steps Out is a 1929 silent comedy-drama film	American Film Institute's list rankin
[5] AFT'S 100 Years100 Stars is a list of the top 25 male and 25 female Hollywood stars The list was unveiled by the American Film	the top 25 male and 25 female greate screen legends of American film histor
Institute on June 15, 1999	and is the second list of the AFI 1
[4] "AFI's 100 Years 100 Stars" Of the 25 actresses listed, Davis	Years series.
was ranked at number two, behind Katharine Hepburn. The United States	
2008 marking the 100th anniversary of her birth	
[5] Joan Crawford was also voted the tenth greatest female star of	
the classic American cinema by the American Film Institute	
Lorenzo Monaco was influenced by which Italian p	ainter and architect?
8 SearChain	🕨 New Bing
Lorenzo Monaco lived and worked between 1370 and 1425 [1]. The	Lorenzo Monaco (c. 1370 - c. 142
tamous Italian painters and architects during his time included Giotto, Masaccio, Brunelleschi, and Donatello [2], Giotto and Masaccio	was an Italian painter of the late Goth
specifically influenced his work [2], as well as Piero di Giovanni in Siena.	Piero di Giovanni in Siena, Italv ar
Spinello Aretino, and Agnolo Gaddi [2]. So, the answer to the question	was apprenticed in Florence. He was
"Lorenzo Monaco was influenced by which Italian painter and architect?"	influenced by Giotto and that artist
is Giotto, Masaccio, Piero di Giovanni in Siena, Spinello Aretino, and	tollowers Spinello Aretino and Agno
renoto statuli.	Guuuti [1].
[1]Lorenzo Monaco Lorenzo Monaco (born Piero di Giovanni; 1370 -	[1]: Lorenzo Monaco Lorenzo Monac
c. 1425) was an Italian painter of the	was an Italian painter of the late Gothi
[2] Lorenzo was an Italian painter of the late Gothic-early Renaissance	early Renaissance age. He was bon
age. He was born riero of Giovanni in Siena. Little is known about his youth, apart from the fact that he was apprenticed in Florence. He was	known about his vouth, anart from th
influenced by Giotto and his followers Spinello Aretino and Agnolo	fact that he was apprenticed in Florence
Gaddi. In 1390 he joined the Camaldolese monastery of Santa Maria	He was influenced by Giotto and h
degli Angeli. He was thenceforth generally known as "Lorenzo Monaco"	followers Spinello Aretino and Agno
(English: "Lawrence the Monk"). In the 1390s he executed three panels of the Biblioteca Laurenziana for his convent	Gaddi. In 1390 ne joined th Camaldolese
The song Arizona was recorded by Paul Revere and Mark Lin	dsay but who wrote the song?
SearChain	New Bing
Paul Revere is Patriot in the American Revolution [1]. Mark Lindsay is a	The song "Arizona" was written I
musician who is the lead singer of Paul Revere & the Raiders [2]. The	Kenny Young and recorded by Ma
song Arizona was recorded by Paul Revere and Mark Lindsay, and	Lindsay, a solo errort while still le
written by Kenny Toung [5]. So the final answer is Kenny Toung.	in 1969 [1]. Is there anything else y
[1] Paul Revere was an American silversmith, engraver, early	would like to know?
industrialist, and Patriot in the American Revolution	[1] U
[2] Mark Lindsay (born March 9, 1942) is an American musician, best	[1]: "Arizona" is a song written Kenny Young and recorded in 1969
[3] "Arizona" is a song written by Kenny Young and recorded in 1969	Mark Lindsay, a solo effort while st
by Mark Lindsay, a solo effort while still lead singer for Paul Revere and	lead singer for Paul Revere and th
the Raiders	Raiders
When was the album with the song Unbelievable by American rap	per The Notorious B.I.G released?
SearChain	New Bing
The Notorious B.I.G., also known as Christopher Wallace, was an	The song "Unbelievable" by America
American rapper [1]. He released two studio albums during his lifetime: "Ready to Die" in 1994 and "Life After Death" in 1997 [2]. The song	rapper ine Notorious B.I.G. w
"Unbelievable" was included on his debut album "Ready to Die	Ready to Die on September 13, 1994 h
[3],"which was released on September 13, 1994 [4]. So the answer is	Bad Boy Records and Arista Record
September 13, 1994 .	[1]. Is there anything else you wou
[1] The Notorious B I G. Christopher Course Laters Walls (24 - 2)	like to know?
1972 – March 9, 1997), known professionally as the Notorious B I G.	[1] Ready to Die   Ready to D
Biggie Smalls, or Biggie, was an American rapper	Ready to Die is the debut studio albu
[2] Life After Death Life After Death is the second and final studio	by American rapper The Notorio
album by American rapper The Notorious B.I.G., released on March 25,	B.I.G., released on September 1
1997 The Notorious B.I.G. turther delving into the matioso rap	1994
[3] Ready to Die Songs on the album range from homicide	
narratives ("Warning") to braggadocios battle rans ("The What "	
narratives ( warning ) to braggadoeios battle raps ( The what,	
"Unbelievable")	
"Unbelievable") [4] Ready to Die   Ready to Die Ready to Die is the debut studio	
<ul> <li>[4] Ready to Die   Ready to Die Ready to Die is the debut studio album by American rapper The Notorious B.I.G., released on September 13, 1994.</li> </ul>	

chain that can solve the complex question, when a sub-question cannot be solved, even without the help of IR, LLM can try to further decompose or rewrite the sub-question to continue reasoning. It is because our method focuses on building a global chain that can solve the complex question (global perspective), rather than answering or generating sub-questions step by step (local perspective). This makes LLM try more when faced with intermediate difficulties to finally solve complex questions. Case study shown in Figure 8  $\sim$ 9 indicates that SearChain continues reasoning but baselines stop.

who followed t	ne father of Menucha Rochel Sionim?
──→ CoT	The question is incomplete and unclear about Menucha Rochel Slonim's father.×
Self-Ask w/o Retriev	Please provide additional information about Menucha Rochel Slonim's father so we can answer the question.×
SearChain w/o Retriev	[Query 1]: Who was the father of Menucha Rochel Slonim? [Answer 1]: Rabbi Dovber Schneuri, the second Rebbe of the Chabad Lubavitch movement. al [Query 2]: Who succeeded Rabbi Dovber Schneuri as Rebbe? [Answer 2]: His son, Rabbi Menachem Mendel Schneersohn. [Final Content]: The father of Menucha Rochel Slonim is Rabbi Dovber Schneuri [1]. His son Rabbi Menachem Mendel Schneersohn succeeded Rabbi Dovber Schneuri as Rebbe[2]. So the final answer is Richard Stallman.
Who is the press Friendship with	ident of the newly declared independent country that is part of the commission of Truth and the country where Inierie is located?
CoT	Unfortunately, there is not enough information provided in the question to accurately answer it.×
Self-Ask w/o Retrieval	Follow up: What is the name of the newly declared independent country? Intermediate answer: The name of the country is not clear. ×
SearChain • w/o Retrieval	[Query 1]: What countries are part of the commission of Truth and Friendship with Inierie? [Answer 1]: East Timor and Indonesia [Query 2]: Which of these countries recently declared independence? [Answer 2]: East Timor [Query 3]: Who is the president of East Timor? [Answer 3]: Francisco Guteres [Final Content]: East Timor and Indonesia are part of the commission of Truth and Friendship with Inierie [1]. East Timor could had name is increasing of the president of it is Francisco Guteres. So the final answer is increasing of the president of it is
Where did Willi	am Beckford's child die?
	William Beckford's child, a daughter named Maria Hamilton, the information about the death of Maria Hamilton is not clear $\times$
Self-Ask w/o Retrieval	Follow up: What was the name of William Beckford's child? Intermediate answer: The name of William Beckford's child is not specified in the question. ×
SearChain w/o Retrieval	[Query 1]: Who is William Beckford? [Answer 2]: William Beckford was an English novelist, art collector, and politician. [Query 2]: Did William Beckford have any children? [Answer 2]: Yes, William Thomas Beckford have any children? [Answer 3]: William Thomas Beckford live? [Answer 3]: William Thomas Beckford live? [Answer 3]: William Thomas Beckford's child die? [Query 4]: Where did William Beckford's child die? [Answer 4]: William Thomas Beckford's child die? [Answer 4]: William Thomas Beckford's child die? [Inal Content]: William Beckford's as on named William Thomas Beckford [2]. William Thomas Beckford lived in various places in English novelist, art collector, and politician [1]. William Beckford had one child, as on named William Thomas Beckford [2]. William Thomas Beckford lived in various places in England, including Bath and Fonthill Abbey and died in Bath [3,4]. So the final answer is Bath. ∽
Who is the own	er of the record label of the performer of Groovy Little Summer Song?
→ CoT	The owner of the record label of the performer of "Groovy Little Summer Song" is not specified. More information is needed to answer this question.×
→ Self-Ask w/o Retrieval	Follow up: Who is the performer of Groovy Little Summer Song? Intermediate answer: The performer of Groovy Little Summer Song is not specified.×
→ SearChain w/o Retrieval	[Query 1]: What is the name of the performer of Groovy Little Summer Song? [Answer 1]: I don't know. [Query 2]: Can you find out the name of the performer of Groovy Little Summer Song? [Answer 2]: The performer of Groovy Little Summer Song is James Otto. [Query 3]: What is the name of the record label that James Otto is signed to? [Answer 3]: Warner Music Nashville is the record label that James Otto is signed to? [Answer 4]: Warner Music Nashville is a subsidiary of Warner Music Group, so the owner would be Warner Music Group. [Final Content]: The performer of Groovy Little Summer Song is James Otto [2]. Warner Music Nashville is the record label that James Otto is signed to [3]. The owner of Warner Music Nashville is the record label that James Otto is signed to [3]. The owner of Warner Music Nashville is Warner Music Group [4]. So the final answer is Warner Music Group, <sub>&gt;&gt;</sub>

Figure 8: Case study for CoQ vs Baselines in Reasoning.

A.1.2 Case Study for CoQ vs Baselines in Reasoning. Baselines focus on solving local sub-questions while ignoring the global plan-ning of the reasoning chain, which leads LLM to tend to stop reason-ing rather than try more when a sub-question cannot be solved. In our method, LLM acts as a commander that plans a global reasoning

A.2.1 Knowledge-reasoning Ability Affected by the Existing Form of Knowledge. We divide the knowledge of LLM into two forms of existence, one exists in the parameters, and the other exists in the input prompt. In this section, we explore the effect of two different forms on the knowledge-reasoning ability of LLM. The experiments are performed on *gpt-3.5-turbo*.

Table 7: Accuracy on Sall

	2-Hop	3-Нор	4-Hop
Accuracy	0.394	0.317	0.093

Specifically, we obtain the sub-questions corresponding to each complex question from the datasets provided by Musique, and then we select the complex questions ( $\mathbb{S}_{all}$ ) that LLM can answer all subquestions of them from full datasets. Table 7 shows the accuracy of LLM in answering the complex questions in  $S_{all}$ , which indicates that even though the parameters have memorized the answer to each sub-question of the complex question, LLM cannot compose these answers to effectively reason the complex questions.

Furthermore, we explore the effect of the form of knowledge on reasoning ability. Specifically, we select the complex questions  $(S_i)$ that LLM can answer the sub-questions for the last *i* steps. And input the answers of sub-questions for the first n - i steps (n is the number of sub-questions) in the form of prompt. In this way, when LLM reasoning the complex questions, the knowledge of n - i sub-questions comes from the prompt, and the knowledge of i sub-questions comes from the parameters. Figure 10(a) shows the performance varies with *i*, which indicates that when *i* increases, the knowledge in prompt decreases, the knowledge in parameters increases, and the accuracy of LLM decreases. To prevent the impact of the leakage of sub-questions in the prompt, we also conduct the same experiment on  $\mathbb{S}_{all},$  except that the prompt only contains subquestions without the answers (knowledge is still from parameters). The results are shown in Figure 10(b). The accuracy of Figure 10(b) is lower than that in Figure 10(a), which further confirms that the LLM has the stronger knowledge-reasoning ability for complex questions when the knowledge exists in the prompt and shows that the paradigm of using IR to provide knowledge for LLM can not only help LLM acquire its unknown knowledge but also improve the knowledge-reasoning ability of LLM.

A.2.2 Conflict in the Knowledge of IR and LLM. Although IR can provide additional knowledge to LLM, contradictory knowledge in IR and LLM can negatively affect the performance of LLM. We use five Open-domain question-answering datasets (Natural Questions, TriviaQA, WebQuestions, SQuAD, and HotpotQA) to detect the knowledge in LLM and IR. We select the set of questions that LLM gives correct answers ( $S_t$ ) and the set that LLM gives wrong answers ( $\mathbb{S}_f$ ). We evaluate the probability that IR can rank the correct document at Top-1 on  $\mathbb{S}_t$  and  $\mathbb{S}_f$  respectively. The results shown in Table 8 indicate that although IR can provide LLM with its unknown or wrong knowledge (Top-1 on  $S_f$ ), it can also interfere with correct knowledge in LLM because Top-1 on  $S_t$  only reach (60%  $\sim$  73%) and accuracy of LLM drops from 100% to (80%  $\sim$  90%). This indicates that decoupling the knowledge of LLM and IR in

→ CoT	The question is unclear and may need more context. Please provide more information. $\times$
Self-Ask	Follow up: Which President are we referring to? Intermediate answer: The question does not provide a specific President, so we would need
	additional information to answer it accurately. Follow up: Which body are we referring to?
	Intermediate answer: The question refers to a specific body that the President
	Without additional information, we cannot answer the question accurately.×
SearChain	[Query 1]: What is the body the President calls on for support of his appointments to the USAE?
w/o Retrieval	[Answer 1]: The Senate Armed Services Committee.
	[Query 2]: Which party had a majority in the Senate Armed Services Committee during the period in question?
	[Answer 2]: The Republican Party.
	[Query 3]: When did the Republican Party gain control of the House of Representatives? [Answer 3]: In the 2010 elections
	[Final Content]: the President calls on the Senate Armed Services Committee for support of
	appointments to the USAF [1]. The Republican Party had a majority in the Senate Armed
	Representatives in the 2010 elections [3]. So the final answer is the 2010 elections.
When did the bi	rthplace of Ah Via Musicom's performer become the capital of the state Knowles is from?
	It is unclear who the performer of Ah Via Musicom is and which state Knowles is from.
- C01	Therefore, this question cannot be answered without additional information. ×
	Follow up: Who is the performer of Ah Via Musicom?
Self-Ask	Intermediate answer: The performer of Ah Via Musicom is Eric Johnson. Follow up: Which state is Knowles from?
w/o Retrieval	Intermediate answer: The state Knowles is from is not specified in the original question, so
	this information is missing. Therefore, the final answer cannot be determined. ×
SearChain	[Ouery 1]: Who is the performer of Ab Via Musicom?
w/o Retrieval	[Answer 1]: The performer of Ah Via Musicom is Eric Johnson.
	[Query 2]: What is the birthplace of Eric Johnson?
	[Query 3]: Which state is Knowles from?
	[Answer 3]: Knowles is from the state of Texas.
	[Answer 4]: Austin became the capital of Texas in 1839.
	[Final Content]: The performer of Ah Via Musicom is Eric Joh [1]. Eric Johnson was born
	Austin, Texas [2]. Knowles is from the state of Texa [5]. Austin became the capital of Texa [1839 [4]. So the final answer is 1839. ✓
Where was the p	lace of death of Joseph Ferdinand Of Bavaria's father?
	Joseph Ferdinand Of Bavaria's father was Maximilian II of Bavaria. But the place of deat
+ Col	him is not clear. ×
Self-Ask	Follow up: Who was Joseph Ferdinand Of Bavaria's father?
w/o Retrieva	al Intermediate answer: Joseph Ferdinand Of Bavaria's father was Prince Adalbert of Bavaria Follow up: What was the place of death of Prince Adalbert of Bavaria?
	Intermediate answer: The place of death of Prince Adalbert of Bavaria is not specified. ×
	[Query 1]: Who was Joseph Ferdinand Of Bavaria's father?
SaarCheir	[Answer 1]: Maximilian II Emanuel, Elector of Bavaria. [Ouery 2]: When did Maximilian II Emanuel, Elector of Bavaria die?
w/o Retrieva	[[Answer 2]: February 26, 1726.
	[Query 3]: Where did Maximilian II Emanuel, Elector of Bavaria die? [Answer 3]: Munich Germany
	[Final Content]: Joseph Ferdinand Of Bavaria's father was Maximilian II Emanuel, Elector
	Bavaria [1], Maximilian II Emanuel, Elector of Bavaria died on February 26, 1726 [2].
	Maximilian II Emanuel, Elector of Bayaria died in Munich, Germany [3]. So the answer i
	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany, ~
- What nationality	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother?
- What nationality	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorehe Remenco's mother or her nationali
- What nationality → CoT	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorghe Remenco's mother or her nationalit Follow up: What is the nationality of Gheorghe Remenco'?
- What nationality → CoT [ → Self-Ask states of the second sec	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorghe Remenco's mother or her nationali Follow up: What is the nationality of Gheorghe Remenco' Interrediate answer: The nationality of Gheorghe Remenco is Romanian.
- What nationality → CoT → Self-Ask w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorghe Remenco's mother or her nationali Follow up: What is the nationality of Gheorghe Remenco' Intermediate answer: The nationality of Gheorghe Remenco's father? Follow up: What is the nationality of Gheorghe Remenco's father? Follow up: What is the nationality of Gheorghe Remenco's father?
- What nationality → CoT [ → Self-Ask w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorghe Remenco's mother or her nationali Follow up: What is the nationality of Gheorghe Remenco' Intermediate answer: The nationality of Gheorghe Remenco's father? Intermediate answer: The is not known what the nationality of Gheorghe Remenco's father? Intermediate answer: The nationality of Gheorghe Remenco's father?
- What nationality → CoT [ → Self-Ask w/o Retrieval SearChain [	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorghe Remenco's mother or her nationali Follow up: What is the nationality of Gheorghe Remenco' Intermediate answer: The nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father? [Query 1]: Who is Gheorghe Remenco? [Answer 1]: There is not enough information to answer this question.
- What nationality → CoT [ → Self-Ask w/o Retrieval → SearChain w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~           is Gheorghe Remenco's mother?           There is no information readily available about Gheorghe Remenco's mother or her nationalit           Follow up: What is the nationality of Gheorghe Remenco's father?           Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father?           Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father?           [Query 1]: Who is Gheorghe Romenco?           [Answer 1]: There is not enough information to answer this question.           [Query 2]: Can we find any information about Gheorghe Remenco?
- What nationality → CoT → Self-Ask w/o Retrieval → SearChain w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~           is Gheorghe Remenco's mother?           There is no information readily available about Gheorghe Remenco's mother or her nationalit           Follow up: What is the nationality of Gheorghe Remenco's father?           Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father?           Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father?           Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father?           Query 1]: Who is Gheorghe Remenco?           [Answer 1]: There is not enough information to answer this question.           Query 2]: Can we find any information about Gheorghe Remenco?           [Answer 2]: Yes, Gheorghe Remenco was a Romanian astronaut who died in the 1989           Romanian Revolution.
- What nationality → CoT → Self-Ask w/o Retrieval → SearChain w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~           is Gheorghe Remenco's mother?           There is no information readily available about Gheorghe Remenco's mother or her nationalit Follow up: What is the nationality of Gheorghe Remenco? Intermediate answer: The nationality of Gheorghe Remenco's father? Intermediate answer: The nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father was [Query 1]: Who is Gheorghe Remenco? [Answer 1]: There is not enough information to answer this question. [Query 2]: Yes, Gheorghe Remenco was a Romanian astronaut who died in the 1989 Romanian Revolution. [Query 3]: Is there any information about Gheorghe Remenco's family?
- What nationality → CoT [ → Self-Ask w/o Retrieval → SearChain w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~           is Gheorghe Remenco's mother?           There is no information readily available about Gheorghe Remenco's mother or her nationalit Follow up: What is the nationality of Gheorghe Remenco's father? Intermediate answer: The nationality of Gheorghe Remenco's father? Intermediate answer: The nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father was [Query 1]: Who is Gheorghe Remenco? [Answer 1]: There is not enough information to answer this question. [Query 2]: was of fadary information about Gheorghe Remenco? Answer 2]: Yes, Gheorghe Remenco was a Romanian astronaut who died in the 1989 Romanian Revolution. [Query 3]: Yes, Gheorghe Remenco & father was Russian. Final Content]: Gheorghe Remenco was a Romanian, while his mother was Russian.
- What nationality → CoT [ → Self-Ask w/o Retrieval → SearChain w/o Retrieval	Maximilian II Emanuel, Elector of Bavaria died in Munich, Germany [3]. So the answer i Munich, Germany. ~ is Gheorghe Remenco's mother? There is no information readily available about Gheorghe Remenco's mother or her nationali Follow up: What is the nationality of Gheorghe Remenco's father? Intermediate answer: It is not known what the nationality of Gheorghe Remenco's father? (Answer 1]: Who is Gheorghe Remenco? [Answer 1]: There is not enough information about Gheorghe Remenco? [Answer 2]: Yes, Gheorghe Remenco was Aromanian astronaut who died in the 1989 Romanian Revolution. [Final Content]: Gheorghe Remenco was a Romanian stronaut who died in the 1989 Romanian Revolution [2]. His father was Romanian, while his mother was Russian [3]. Therefore,

Figure 9: Case study for CoQ vs Baselines in Reasoning.

# A.2 Analysis of IR for LLM in Complex **Knowledge-Intensive Tasks**

In this section, we analyze the effect of IR on LLM on complex knowledge-intensive tasks from perspectives of knowledge-reasoning ability and conflict in the knowledge of IR and LLM.

1393Table 8: Top-1 on  $S_t$  and  $S_f$ . IR model is ColBERTv2. Red means worse performance after adding IR, green means better1394performance after adding IR.

	$NQ_{S_t}$	$NQ_{S_f}$	$WQ_{S_t}$	$WQ_{S_f}$	$\operatorname{Trivia}_{\mathbb{S}_t}$	$\operatorname{Trivia}_{\mathbb{S}_f}$	$Squad_{S_t}$	$\operatorname{Squad}_{\mathbb{S}_f}$	$HotpotQA_{S_t}$	$\operatorname{HotpotQA}_{\mathbb{S}_f}$
IR Top-1	60.99	29.20	63.24	26.07	72.60	28.10	65.88	40.67	55.61	29.13
LLM Acc.	100	0	100	0	100	0	100	0	100	0
LLM w/ IR Acc.	80.75	25.43	85.34	23.17	89.47	25.32	86.45	30.42	80.42	19.86

Table 9: Performance of SearChain and DSP on complex knowledge-intensive tasks on Vicuna-13B. Bold denotes the best result in different settings. FC: Fact Checking, LFQA: Long-Form QA. Metric for LFQA: ROUGE-L. Metric for others: cover-EM.

		Muti-H	Iop QA		Slot	Filling	FC	LFQA
	HoPo	MQ	WQA	SQA	zsRE	T-REx	FEV.	ELI5
		Intera	ction with	ı Informati	ion Retriev	val		
DSP	25.45	9.06	27.50	62.01	33.71	49.08	73.05	22.58
SearChain	29.77	10.59	32.32	63.75	36.86	52.75	75.47	24.05



Figure 10: Accuracy varies with knowledge form. The larger *i* is, the larger the proportion of knowledge comes from parameters, and the smaller the proportion of knowledge comes from prompt.

an explainable way that only provides LLM with its unknown or wrong knowledge to avoid misleading LLM is important.

## A.3 Performance on Vicuna-13B

In this section, we compare SearChain with the competitive baseline DSP on Vicuna-13B<sup>6</sup>, a strong open source large model <sup>7</sup> trained by Stanford. The experimental results in Table 9 show that SearChain again outperforms DSP on Vicuna-13B.

# A.4 Experimental Details

*A.4.1* Threshold Selection. As for the confidence threshold  $(\theta)$ , we initialize the initial value of the confidence threshold (1.0) based on prior knowledge and gradually increase the value with a step size of 0.1. We validate the F1-score (a comprehensive metric of the Recall and Precision of judging whether the passage can answer the question) on the mixed open-domain QA datasets (NQ, TriviaQA, WebQ, and TREC) after each value change. We find that when the

<sup>6</sup>https://lmsys.org/blog/2023-03-30-vicuna/

### 

# Table 10: Performance change with ROUGE threshold.

	$\alpha = 0.30$	$\alpha = 0.35$	$\alpha = 0.40$	$\alpha = 0.45$	$\alpha = 0.50$
Performance	25.50	25.57	25.58	25.57	25.55

confidence threshold is 1.5, the highest F1-score can be achieved so we set the confidence threshold as 1.5. As for the ROUGE threshold ( $\alpha$ ), we determine this value by manually observing the ROUGE relationship between the generated text and the ground truth in the few examples in in-context learning. Our further experiments in Table 10 show that when the value range of ROUGE threshold is between 0.3 and 0.5, the performance change on ELI5 is not obvious.

*A.4.2 Number of Examples in Prompt.* We show the number of examples in prompt used for in-context learning on different datasets (Table 11). Our method (SearChain) achieves the best performance with fewer learning examples than competitive baselines.

Construct a global reasoning chain for this complex question [Question]:" {}" and answer the question, and generate a query to the search engine based on what you already know at each step of the reasoning chain, starting with [Ouerv].
You should generate the answer for each [Query], starting with [Answer].
You should generate the final answer for the [Question] by referring the [Query]-[Answer] pairs, starting with [Final Content].
For exmaple:
[Question]:"How many places of higher learning are in the city where the Yongle emperor greeted the person to whom the edict was addressed?"
[Query 1]: Who was the edict addressed to?
[Answer 1]: the Karmapa
[Query 2]: Where did the Yongle Emperor greet the Karmapa?
[Answer 2]: Nanjing
[Query 3]: How many places of higher learning are in Nanjing?
[Answer 3]: 75
[Final Content]: The edict was addressed to Karmapa [1]. Yongle Emperor greet the Karampa in Nanjing [2]. There are 75 places
of higher learning are in Nanjing [3]. So the final answer is 75.
[Ouestion]:"Which magazine was started first Arthur's Magazine or First for Women?"
[Query 1]: When was Arthur's Magazine started?
[Answer 1]: 1844.
[Query 2]: When was First for Women started?
[Answer 2]: 1989
[Final Content]: Arthur's Magazine started in 1844 [1]. First for Women started in 1989 [2]. So Arthur's Magazine was started
first. So the final answer is Arthur's Magazine.

# Figure 11: Prompt for generating Chain-of-Query on HotpotQA, Musique, WikiMultiHopQA, zsRE and T-REx (in the setting without information retrieval).

A.4.3 Prompts in Experiment. We show the prompt used in experiment on different datasets in Figure  $11 \sim 13$ .

<sup>1449 &</sup>lt;sup>7</sup>https://huggingface.co/lmsys/vicuna-13b-delta-v1.1/tree/main

# Table 11: Number of examples in prompt used for in-content learning on different datasets.

	N	luti-H	Hop QA	ł	Slot	Filling	FC	LFQA
	HoPo	MQ	WQA	SQA	zsRE	T-REx	FEV.	ELI5
W	ithout	Infor	matior	ı Retri	ieval			
Direct Prompting	0	0	0	0	0	0	0	0
Auto-CoT	4	4	4	6	4	4	4	2
CoT	4	4	4	6	4	4	4	2
CoT-SC	4	4	4	6	4	4	4	2
Recite-and-answer	4	4	4	6	4	4	4	2
Self-Ask w/o IR	4	4	4	6	4	4	4	2
Least-to-Most	4	4	4	6	4	4	4	2
Plan-and-Solve	4	4	4	6	4	4	4	2
SearChain w/o IR	2	2	2	6	2	2	4	2
Interac	ction w	vith I	nforma	tion I	Retriev	val – – –		
Direct Retrieval	0	0	0	0	0	0	0	0
ToolFormer	4	4	4	6	4	4	4	2
Self-Ask	4	4	4	6	4	4	4	2
Plan-and-Solve w/ IR	4	4	4	6	4	4	4	2
React $\rightarrow$ CoT-SC	6	4	4	6	4	4	4	2
Verify-and-Edit	2	2	2	2	2	2	4	2
Tree-of-Thought w/ IR	4	4	4	6	4	4	4	2
DSP	16	8	8	8	8	8	8	2
SearChain	2	2	2	2	2	2	4	2
Construct a global reasoning chain for this the search engine based on what you alrea You should generate the answer for each [	s complex o idy know at Query], sta	uestion each ste rting wit	[Question]: p of the rea h [Answer]	"{}" and a soning cha	nswer the ain, startin	question, and g with [Quer	d generate y].	a query to
Construct a global reasoning chain for this the search engine based on what you alrea You should generate the answer for each You should generate the final answer for t For exmaple:	s complex of dy know at Query], sta he [Questic	uestion each ste rting wit m] by re!	[Question]: p of the rea h [Answer] ferring the [	"{}" and a soning cha Query]-[A	nswer the ain, startin .nswer] pa	question, and g with [Quer irs, starting v	d generate y]. vith [Fina	e a query to 1 Content].
Construct a global reasoning chain for this the search engine based on what you alrea You should generate the answer for each You should generate the final answer for t For exmaple: [Question]: What causes the trail behind, [Query 1]: What is the trail behind, given at	s complex o dy know at Query], sta he [Questic jets at high high altitud	uestion each ste rting wit on] by res altitude?	[Question]: p of the rea h [Answer] ferring the [ ?"	"{}" and a soning cha Query]-[A	nswer the ain, startin .nswer] pa	question, an g with [Quer irs, starting v	d generate y]. vith [Fina	: a query to l Content].
Construct a global reasoning chain for thii the search engine based on what you alrea You should generate the answer for each You should generate the final answer for t For exmaple: [Quest10:]."What is the trail behind jets at [Answer 1]: The trail behind jets at high a formed when the hot exhaust gases from	s complex o idy know at Query], sta he [Questic jets at high high altitud litude is co n a jet engi	uestion each ste rting wit on] by rel altitude? de? ommonly ne mix w	[Question]: p of the rea h [Answer] ferring the [ m referred to ith the cold	"{}" and a soning cha Query]-[A as a contr er air at hi	nswer the ain, startin .nswer] pa ail, which gh altitude	question, and g with [Quer irs, starting v is short for " 25, causing th	d generate y]. with [Fina condensa ne water v	a query to l Content]. tion trail." apor in the
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each You should generate the final answer for I For exmaple: [Query 1]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a formed when the hot exhaust gases from to condense and freeze into inty ice crysta [Duery 2]: Why do the hot exhaust or asses:	s complex of dy know at Query], sta he [Questic jets at high high altitud lititude is co n a jet engii als. mix with th	uestion each ste rting wit n] by rel altitude? mmonly ne mix w e colder	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a	"{}" and a soning cha Query]-[A as a contr er air at hi ltitudes?	nswer the ain, startin .nswer] pa ail, which gh altitude	question, and g with [Quer irs, starting v is short for " es, causing th	d generate y]. with [Fina condensa ae water v	a query to l Content]. tion trail." apor in the
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each You should generate the final answer for I For exmaple: [Query 1]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a 5 formed when the hot exhaust gases from to condense and freeze into inty ice crysta [Answer 2]: The hot exhaust gases from a between the engine exhaust and the surrow	s complex c dy know at Query], sta he [Questic jets at high high altitud lititude is cc n a jet engin ls. mix with th jet engine : moding air c	uestion   each ste rting wit n] by re! altitude? de? ommonly ne mix w e colder mix with At high a	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a the colder litudes the	"{}" and a soning cha Query]-[A as a contr er air at hi lititudes? air at high	nswer the ain, startin .nswer] pa ail, which gh altitude altitudes o	question, and g with [Quer irs, starting v is short for " es, causing th due to the pro-	d generate y]. with [Fina condensa te water v essure diff	a query to l Content]. tion trail." apor in the ference
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each You should generate the final answer for I For extnaple: [Query 1]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at [Answer 1]: The trail behind jets at [Answer 2]: The trail behind jets at general weak of the second second second to condense and freeze into imy ice crysta [Answer 2]: The hot exhaust gases from between the engine exhaust and the surrot exhaust gases to rapidly expand and cool, [Ourery 3]: Why done the outer surves in the second se	s complex of dy know at Query], sta he [Questic jets at high high altitude is con a jet engine ink with th jet engine : anding air. / creating a l	uestion each ste rting with o] by ref altitude? de? ommonly ne mix w e colder mix with At high a turbulent mase and	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a the colder : lititudes, the wake behin freeze inte	"{}" and a soning cha Query]-[A as a contr er air at hi lititudes? air is muc d the airc is corter	nswer the ain, startin .nswer] pa ail, which gh altitude altitudes o ch colder a raft. s?	question, and g with [Quer irs, starting v is short for " es, causing th due to the pro- nd thinner, v	d generate y]. with [Fina condensa ae water v essure diff which cau:	a query to l Content]. tion trail." apor in the ference ses the
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each You should generate the final answer for I For exmaple: [Question]: "What causes the trail behind [Query 1]: The trail behind jets at [Answer 1]: The trail behind jets at [Answer 1]: The trail behind jets at [Answer 2]: The trail behind jets at [Answer 2]: The bot exhaust gases from to condense and freeze into inty ice crysts [Answer 2]: The hot exhaust gases from exhaust gases to rapidly expand and cool, causer 3]: Why does the water vapor in it [Query 2]: Why does the water vapor in the envert should be the second second second second second [Query 2]: The water vapor in the size	s complex c dy know at Query), stat he [Questic jets at high high altituu lititude is cc n a jet engine is. mix with th jet engine inding air. A creating a l he air conde dendenses an constitution	uestion each ste rting wit n] by ref altitude? de? mmonly a mix with At high a turbulent ms and l freezes heart of	[Question]: p of the rea h [Answer] ferring the [ " " referred to ith the cold air at high a the colder ititudes, the wake behin freeze into i into ice cry o high term	"{}" and a soning chi Query]-[A as a contr er air at high air at high air is mud dithe airc ice crystal: stals becau	nswer the ain, startin unswer] pa ail, which gh altitudes o ch colder a raft. s? use the excl	question, an g with [Quer irs, starting v is short for " es, causing th due to the pro- nd thinner, v naust gases fi	d generate y]. with [Fina condensa ie water v essure diff which cause room the je	a query to l Content]. tion trail." 1 apor in the ierence ses the t engine
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for t For extnaple: [Query 1]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a formed when the hot exhaust gases from to condense and freeze into inny ice crysta [Query 2]: Why do the hot exhaust gases from between the engine exhaust and the surrou exhaust gases to rapidly expand and cool, [Query 3]: Why does the water vapor in the Answer 3]: The water vapor in the air con contain a significant amount of water vapor with the cold air at high altimades, it quick	s complex of dy know at Query], sta he [Questic litude is co n a jet engine : ls. mix with th jet engine : anding air. 2 creating a t e air conde dor, which is ly loses heat	uestion each ste rting wit n] by rel altitude? de? mmonly he mix with At high a turbulent mse and l freezes heated t	[Question]: p of the rea h [Answer] erring the [ " referred to ith the cold air at high a the colder : litudes, the wake behir freeze into i into ice cry o high temp ergy, causin	"{}" and a soning chu Query]-[A as a contr er air at hi air at high air is mu id the airc: air air air air air is mu d the airce rystal stals became reatures in g the wate	nswer the ain, startin ail, which gh altitudes of the colder a raft. s? use the exl n the engir r vapor to	question, and g with [Quer irs, starting v is short for " es, causing the due to the pre- nd thinner, v naust gases fi e.e. When the condense ini	d generate y]. with [Fina condensa ie water v ssure diff which caus rom the je hot water to liquid d	a query to l Content]. tion trail." 1 apor in the ference ses the t engine vapor mix roplets and
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for t For extnaple: [Query 1]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a is formed when the hot exhaust gases from to condense and freeze into inny ice crysts [Answer 2]: The hot exhaust gases from a between the engine exhaust and the surroo exhaust gases to rapidly expand and cool, [Query 2]: Why do the hot exhaust gases (Answer 2]: The hot exhaust gase from a significant amount of water vapor in th any or in the air cor- contain a significant amount of water vapor with the cold air at high altitudes, it quick then freeze into ice crystals.	s complex c dy know at Query], sta Query], sta Power of the description of the description description of the description of th	uestion each ste rting wit on] by rel altitude de? mmonly ne mix w e colder mix with At high a urrbulent ense and l freezes heated t it and en-	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a the colder. Ititudes, the wake behin freeze into into ice cry o high temp ergy, causin d when the l	"{}" and a soning chi Query]-[A as a contr er air at high air at high air is mu d the airc: air at high air is mu d the airce rystal stals becan seratures in g the wate hot exhaus	nswer the ain, startin answer] pa ail, which gh altitudes of ch colder a raft. se the exp n the engir r vapor to t gases fro	question, and g with [Quer irs, starting v is short for " es, causing th due to the pro- nd thinner, v aust gases fi e. When the condense in m a jet englin	d generate y]. with [Fina condensa te water v essure diff which cau: rom the je hot water to liquid d ne mix wi	a query to l Content]. tion trail."   apor in the erence ses the t engine vapor mix troplets and th the coldo
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for t For example: [Quesrtion]: What causes the trail behind [Quesr 1]: The trail behind jets at [Answer 1]: The trail behind jets at [Answer 2]: The hot exhaust gases from to condense and freeze into tiny ice crysts (Answer 2]: The hot exhaust gases from a between the engine exhaust and the surror exhaust gases to rapidly expand and cool, [Quesr 2]: Why do the hot exhaust gases (Answer 2]: The hot exhaust gases from a between the engine exhaust and the surror contain a significant amount of water vapor in the air at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hij air at high altitudes, causing the water vapor with the cold ar in due to the pressure diff	s complex c dy know at Query], sta he [Questic jets at high high altitude is contained and the state mix with th jet engine : anding air. 2, creating a 1 he air conde dnenses and creating a the endenses and gh altitude y loses hea gh altitude ber in the air conde dnenses and core in the air conde dnenses and core in the air conde dnenses hear conde gh altitude ber in the state of the state core in the air conde dnense hear conde dnenses hear conde dnenses and core in the air conde dnense hear conde dnenses hear conde dne	uestion each stern] by rel altitude? de? ommonly ne mix with At high a turbulent ense and l freezes heated t t and en- is formed veen the	[Question]: p of the rea h [Answer] ferring the [ "" referred to ith the cold air at high a the colder . litudes, the wake behin freeze into is into ice cry o high temp ergy, causin d when the l ense and free ense and free ngine exha	"{}" and a soning cha Query]-[A as a contr er air at hi ditudes? air at high air is mud d the airc ice crystal stals becan eratures in g the wate not exhaus ecze into t sub expension.	nswer the ain, startin .nswer] pa ail, which gh altitudes of ch colder a raft. sse the exl sse the exl n the engir r vapor to at gases fro iny ice cry	question, an g with [Quer irs, starting v is short for " es, causing the due to the pro- nd thinner, v haust gases fi e.e. When the condense in ma jet engit stals (1]. The ling air (2), a	d generate y]. with [Fina condensa te water v essure diff which cause rom the je hot water to liquid d the exhaust und the water	a query to I Content]. tion trail." apor in the erence ses the t engine vapor mix troplets and th the cold gases mix ter vapor
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for to for extnaple: [Question]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a fanswer 2]: The hot exhaust gases from to condense and freeze into imy ice crysta (Answer 2]: The hot exhaust gases from a between the engine exhaust and the surroo to condense and freeze into imy ice crysta (Answer 2]: The hot exhaust gases from a between the engine exhaust and the surroo contain a significant amount of water vapor in it (Answer 3]: Why does the water vapor in the air co contain a significant amount of water vapor with the cold er at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hij air at high altitudes, causing the water vapor in with condenses and freezes because it is heated (ondenses and freezes because it is heated with the cold er at high altitudes [3].	s complex of dy know at Query], sta he [Questic jets at high high altituc litude is co na jet engine : anding air. / creating a 1 with th jet engine : anding air. / creating a 1 with with th jet second expension of the second or, which is ly loses hea gh altitude or in the ai creace betw to high ter	uestion 1 each ste itring with more all each ste each ste	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a it at high a it at high a it at high a the colder. Ititudes, the wake behin freeze into into ice cry o high temp grgy, causin d when the l ense and fr engine exha s in the eng	"{}" and a soning chi Query]-[A as a contr er air at hi ditudes? air at high air at high a	nswer the ain, startin .nswer] pa ail, which gh altitudes of ch colder a raft. se the exl se the exl the exl t	question, an g with [Quer irs, starting v is short for " es, causing the due to the pro- nd thinner, v saust gases fi e.e. When the condense ini om a jet engin stals [1]. The ting air [2], <i>k</i> vlass shals (1).	d generate y]. with [Fina condensa te water v essure diff which cau: rom the je hot water to liquid d ne mix wi e exhaust und the ws und energy	a query to I Content]. tion trail." 1 apor in the ference ses the t engine vapor mix iroplets and th the cold gases mix uter vapor
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for to for exmaple: [Query 1]: What causes the trail behind (Query 1]: That is the trail behind jets at (Answer 2]: The trail behind jets at high a Answer 2]: The trail behind jets at high a Answer 2]: The tot exhaust gases from to condense and freeze into imy ice crysta (Answer 2]: The hot exhaust gases for to condense and freeze into imy ice crysta exhaust gases to rapidly expand and cool, (Query 3]: Why does the water vapor in the air cor- contain a significant amount of water vapor with the cold ar at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hij air at high altitudes, causing the water vapor in the with the cold ar in due to the pressure diff condenses and freezes because it is heated with the cold ar in thigh altitudes [3]. [Question]: In Trading Places (1983, Akre	s complex c dy know at Query], sta he [Questic left and the loguestic person and the loguestic liftude is cc in a jet engine : anding air. J anding air. J andir. J anding air. J anding air. J anding air. J anding air. J anding	uestion in each ste altitude? de? mmnonly ae mix w e colder mix with thigh a urbulent mix with thigh a urbulent is formere to conder to	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a the colder. Ititudes, the wake behin freeze into i air at high a the colder. Ititudes, the wake behin freeze into i into ice cry o high temp rgy, causin d when the I ense and fr engine exht es in the eng	"{}" and a soning cha Query]-[A as a contr er air at hid ditudes? air at high air is mu de the airce istals becau eratures in g the wate hot exhaus ezez into t uust and th gine and the me at the	nswer the ain, startin unswer] pa ail, which gh altitudes of the colder a raft. s? use the expire the expiret the ex	question, and g with [Quer irs, starting v is short for " es, causing th due to the pre- naust gases fi cause of the pre- ondense into manual state of the pre- condense into manual state of the pre- condense into the pre- condense into the pre- condense into the pre- set of the pre- condense into the pre- the pre- condense into the pre	I generate y]. with [Fina condensa condensa e water v sesure diff which caue sesure diff which caue thot water thot water	a query to l Content]. tion trail." 1 apor in the cerence ses the t engine vapor mix troplets and th the cold gases mix ter vapor y when mix suld buying
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for to for example: [Quesrion]: What causes the trail behind (Query 1]: What is the trail behind jets at Answer 1]: The trail behind jets at high a (Answer 2]: The trail behind jets at (Answer 2]: The bot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The bot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The bot exhaust gases from a between the engine exhaust and the surroo with the cold air at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at high air at high altitudes, causing the water vapor in the with the cold air at high altitudes [3]. [Question]: In Trading Places (1983, Akre to for QI at a high price ruin the Duke Bre	s complex c dy know at Query], sta he [Quest], sta he [Quest] jets at high high altituc liftude is cc na a jet engin ls. mix with th jet engine : anding air. Julian and the sta e air conde dnenses and creating a 1 denses and or in the ai erence betv l to high ter byd/Murphy thers? o f Tradine 0 f Tradine 0	question i each stead riting with n] by rel le? mmmonly ac mix with At high a with At high a heated i freezess to condu- te and en- is formere to condu- sing formere to condu- to condu- condu- to condu- to condu- to condu- to condu- to condu- to	[Question]: p of the rea h [Answer] ferring the [ " referred to ith the cold air at high a air at high a the colder h the col	"{}" and a Query]-[A query]-[A as a contr ar at high diftades? ar at high air is muu diftades? g the wate are arystal at a becar aris muu g the wate not exhaus zerz into t suust and th me at the o	nswer the ail, startin ail, which gh altitudes ( h colder a set he colder set he colde	question, and g with [Quer irs, starting v is short for " s, causing th due to the pre- naust gases fi cause of the pre- naust gases fi cause of the pre- naust gases fi cause of the pre- ondense information stals [1]. The ling air [2], a v loses heat a movie work'	d generate y). condensa condensa e water v essure diffi cause water v essure diffi cause water v essure diffication which cause on the je hot water to liquid to to liquid to to liquid to to liquid to to liquid to the water and the water on the je	a query to l Content]. tion trail." apor in the 'erence ses the ' vapor mix troplets and the cold gases mix ter vapor y when mix uld buying
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for to for example: [Query 1]: What causes the trail behind (Query 1]: What is the trail behind jets at Answer 1]: The trail behind jets at high a (Answer 2]: The hot exhaust gases from to condense and freeze into tiny ice crysta (Query 1]: Why do the hot exhaust gases for to condense and freeze into tiny ice crysta (Answer 2]: The hot exhaust gases from a between the engine exhaust and the surroo contain a significant amount of water vapor in th with the cold ari a high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hij air at high altitudes, causing the water vapor in the with the cold ari due to the pressure diff condenses and freezes because it is heated with the cold ari due to the pressure diff condenses and freezes (1983, Akre of of OJ at a high altitudes [3]. [Query 1]: In trailing Places (1983, Akre of OJ at a high arise at the end Answer 1]: In the movie, the main charac	s complex c dy know at Query], sta he [Questic jets at high high altitut littude is cc an a jet engin ls. mix with th jet engine - e air conde disease and or, which is us ly loses hea gh altitude soor in the ai 'erence betv l to high ter soyd/Murphy thers? of Trading ly R	question 1 each ste rting wit n] by rel de? mmonly he mix w e colder mmonly he mix w e colder mmonly he mix w e colder mmonly he mix w e colder to cond f frezes heated t t and en went he mperature t) how do Places? ay Valer	[Question]: p of the rea h [Answer] erring the [ m referred to ith the cold air at high a the colder : litudes, the wake behin freeze into i nito ice cry o high temp ergy, causin 4 when the l ense and fr ense and fr engine exhi- es in the engine exhi- ses the schee time (Eddie anipulatione	"{}" and a soning chi as a contri- er air at hii liitudes? air at high dif he airc erestatis becar erentures in g the wate to to exhaus g the wate to t	nswer the nin, startin , nswer] pa ail, which gh altitudes ( altitudes ( altit	question, and g with [Quer irs, starting v is short for " es, causing th fue to the pro- nul thinner, v naust gases fi e. When the condense in fing air [2], e y loses heat e movie work' antor course transported or anor i transported or anor i transported or anor i	d generate y). vy). condensa condensa e water v essure diff cau which cau on the je hot water on the je hot water e whaust e exhaust e e	a query to I Content]. tion trail." tion trail." tion trail." transformation to the collection tempore the transformation tempore
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for to For example: [Question]: "What causes the trail behind [Query 1]: That is the trail behind jets at [Answer 1]: The trail behind jets at (Answer 2]: The hot exhauts gases from to condense and freeze into tiny ice crysta (Answer 2]: The hot exhauts gases from the bot exhauts gases from a between the engine exhaust and the surroo exhaust gases to rapidly expand and cool, (Query 3]: Why does the water vapor in the air cot contain a significant amount of water vapor with the cold air at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hij air at high altitudes, causing the water vapor in the with the cold air at high altitudes [3]. [Question]: In Trading Places (1983, Akre to of O1 at a high aptrice (1984, A). [Answer 1]: In the movie, the main charar Aykroyd), execute a plan to bankrupt the Imarket.	s complex we way to a second of the second o	question i each ste tring with n) by ref altitude? mmonly ae mix with te? mmonly ae mix with te? mmonly ae colder aurbulent is formere to cond heated 1 freezes ween the mperature places? ay Valer ers by m he ECCO	Question]: p of the rea frame the flammer of the flammer "" referred to the colder air at high a the colder the colder air at high a the colder fleeze into i tho ies provided the colder fleeze into i high temp regy, causin the end of the colder the colder of the colder so is in the end the colder of the colder the colder of the colder of the colder of the colder the colder of the colder of the colder of the colder the colder of the colder of the colder of the colder of the colder the colder of the colder of	"{}" and a Query]-{A as a contr ilititudes? air at high air is mus if the air is mus of the air control air at high air is mus at a bea air the air a air the air and the air control air at high air is mus at a bea air at high air is mus at a bea air and the air control air at a bea air and the air control air at a bea air and the air air air and air air and air air air and air air air air and air air and air air and air air air air and air air air and air	nswer the nin, startin ail, which alititudes of alititudes	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v naust gases fi e. When the condense in fing air [2], e y loses heat a movie work' artner, Louis atted orange j	I generate y]. y]. with [Fina condensa e water v which cau: rom the je hot water d which cau: rom the je which cau: rom the je rom the je which cau: rom the je rom the je	: a query to I Content]. tion trail." ior trail." ior ni the cerence eses the t engine t vapor mix troplets and the cold gates mix vapor mix troplets and the cold gates mix vapor mix the vapor when mix uld buying e III (Dan
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for the for example: [Query 1]: What causes the trail behind [Query 1]: That is the trail behind jets at Answer 1]: The trail behind jets at high a is formed when the hot exhaust gases from a Query 2]: Why do the hot exhaust gases as Answer 2]: The trail behind jets at Answer 2]: The trail behind jets at Answer 3]: The trail behind jets at Answer 3]: The trail behind jets at Answer 3]: The bot exhaust gases from a between the engine exhaust and the surror valuant gases to rapidly expand and cool, Query 3]: Why does the water vapor in the air cor contain a significant amount of water vapor with the cold er at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at high air at high altitudes, causing the water vapor in the ir at high altitudes, causing the water vapor in with the cold er at high altitudes [3]. [Query 1]: In trailing Places (1983, Akro for of OI at a high apprice ruin the Duke Brc of Answer 1]: In The more site sheat the end (Answer 1]: In the movie, the main charar Aykroyd), execute a plan to bankrupt the 1 market. [Avever 2]: Billy Ray and Louis of thair avever 2]: Billy Ray and Louis of thair aver	s complex is scomplex (query), statistical query), statistical query), the query), statistical query), statistical query, the query of the query of the query of the query is the query of	question in each ste tring with n] by ref le? mmonly action le? mmonly action le? mmonly action action le? mmonly action action le? mmonly action action le? mmonly action le? mmonly action le?	(Question):) p of the treat (I hanswer] erring the [ "" " " referred to it his the colder: " " this the colder " " " " " " " " " " " " " " " " " " "	"{}" and a oning ch Query]-[A as a contri- er air at hi ditudes? air at high air is mu dithe air is mu air is mu air is mu dithe air is mu dithe air is mu dithe air is mu seze into t on to chansus and the air is g the wate er statal stats becar g the wate er statal stats becar g the wate er statal stats becar statal stats becar stats stats becar statstat	nswer the in, startin nswer] pa ail, which gh altitudes ( h colder a altitudes ( h colder a st st st st st st st st st st	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v aust gases fi e. When the condense in mom a jet enging stals [1]. The firm air [2], e v loses heat e movie work? atted orange j ure's upcomit	I generate y]. with [Fina condensa e water v which cau: rom the je hot water v vhich cau: rom the je hot water v which cau: rom the je which cau: rom the je which cau: rom the je which water v which cau: rom the je which water v which cau: rom the je which cau: rom the je rom the j	a query to a l Content]." tion trail." cerence eses the t engine troplets an th the cold gates mix- roplets an th the cold gates mix- evapor mix- the vapor when mix- uld buying an UJ) futures eport, whice
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for the For example: [Question]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a to condense and freeze into tiny ice crysta (Answer 2]: Why do the tot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The tot is bot exhaust gases from the observation of the tot exhaust gases from the condense and freeze into tiny is strained and the surrout of the surrout contain a significant amount of water vapor in the arise that the water vapor in the air cor- contain a significant amount of water vapor with the cold ar at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at high air at high altitudes, causing the water vapor with the cold ari at high altitudes [3]. [Query 1]: Must it he scheme (1933, Akro to of OI at a high apprese (1933, Akro to rol at a high apprese (1933, Akro to rol at a high apprese (1933, Akro to rol at a high apprese (1933, Akro Aykroyd), execute a plan to bankrupt the 1 market. [Query 1]: How do Billy Ray and Louis on functions that a harsh winter has destroyce contracts at a low price before the report i	s complex is scomplex (query), sta- legal query), sta- legal query), sta- legal query), sta- legal query, sta- sta- sta- sta- sta- sta- sta- sta-	question in each ste triting with altitude? de? mimonly de colder mix with e colder mix with e colder mix with e colder mix with freezes heated 1 freezes heated 1 freezes heate	(Question): p of the rear p ferring the [ rearing the [ rearing the [ rearing the colder: rearing the colder: rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colde	"{}" and a oning ch Query]-[A as a contra- er air at hi ditudes? air at high air is mu dithe air is mu air is mu ecce rystal stals becar g the wate ce rystal stals becar g the wate ce rystal stals becar g the wate ce rystal stals becar g the wate ce rystal stals becar g the wate stals of the stals o	nswer the in, startin nswer] pa ail, which gh altitudes (a gh altitudes) gh altitudes gh altitudes altitudes (a start, st st st st st st st st st st	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v aust gases fi e. When the condense in main a jet enging stals [1]. The firm a in [2], e v loses heat e movie work? atted orange j ure's upcomin to purchass price after th	l generatic y]. condensa condensa e water v e water v e water v e water v e water v hich cau rom the je hot water to liquid d e emix wi e whats to liquid d e with v with (Pop Winthorp Winthorp con the i p f FCO J fu FCO J fu FCO J fu e reports	: a query to 1 Content] tion trail crence ses the t engine vapor mix ter vapor when mix uld buying uld buying uld buying crent, whice tures crence, the content of the cold passes mix when mix uld buying rel (Dan support, whice tures crence, whice tures
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for the for example: [Question]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at [Answer 2]: The hot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The to exhaust gases from a between the engine exhaust and the surrow chaust gases to rapidly expand and cool, [Query 3]: Why does the water vapor in the activation of the expansion of the surrow chaust gases to rapidly expand and cool, [Query 3]: Why does the water vapor in the air cor- contain a significant amount of water vapor with the cold ar at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hij air at high altitudes, causing the water vapor in the with the cold er at high altitudes [3]. [Question]: In Trading Places (1983, Akro to of O at a high apprice ruin the Duke Bro (Answer 2]: Billy Ray and Louis of Answer 2]: How do Billy Ray and Louis of Answer 2]: How do Billy Ray and Louis of the market. Answer 2]: Billy Qay and Louis obtain in indicates that a harsh winter has destroyce contracts at a low price before the report i the market has responded to the news of the market has	s complex c dyd know at guery), sa sa he [Query], sa he [Que	question in a set of the set of t	(Question):) po of the rear point ferring the [ rearing the [ rearing the [ rearing the colder: rearing the colder: rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the colder rearing the	"{}" and a Query]-{A as a contr er air at hi lititudes? air is mug d the airc ec erystal d the airc ec erystal	nswer the iin, startin answer] pa ail, which gh altitudes gh altitudes gh altitudes altitudes set the engin r vapor to t gases fir the engin r vapor to t gases for gases of the end of the and his pa concentration of Agricult Information	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v aust gases fi e. When the condense in more a set enging stals [1]. The firm air [2], e v loses heat e movie work? atted or ange j ure's upcomin to purchase price after th	l generati y]. condensa condensa e water v essure diffi cau condensa with cau condensa e extra v e exhaust e exhaust e exhaust winthore e exhaust winthore e exhaust winthore e exhaust condensa e exhaust	: a query te 1 Content]. 1 Content]. 1 Content]. 1 Content]. 2 Content 2 Con
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for the for example: [Question]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at sigh a is formed when the hot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The to texhaust gases from a between the engine exhaust and the surrou- exhaust gases to rapidly expand and cool, [Query 3]: Why does the water vapor in the activation of the air con- contain a significant amount of water vapor with the cold ar at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hig air at high altitudes, causing the water vapor in the with the colder air due to the pressure diff condenses and freezes because it is benetic with the cold exist in the shirt of estart Aykroyd), execute a plan to bank rupt the I Answer 1]: In the movie, the main chara. Aykroyd), execute a plan to bankrupt the I andet: [Query 1]: How do Billy Ray and Louis an [Answer 2]: How dos Billy Ray and Louis of the market. [Query 3]: Why does buying a lot of OJ a [Answer 3]: The Duke Brothers, who are; [Answer 3]: The Duke Brothers, who are;	s complex c s complex dy know at g query), sats the query of the query is the query of the query of the query is the query of the query of the query is the query of the query	question in each ste riring with altitude? fe? ownmonly o	(Question):) p of the reis p of the reis p of the reis ferring the [ "" referred to it the colder: "" this the colder it the colder it the colder in the colder it the colder it the colder of high temp regy, causing when the list of the colder of high temp regy, causing when the list of the colder in the colder it the colder is the colder of high temp regy, causing when the list of the colder is the colder	"{}" and a Query]-{A as a contr er air at hi litudes? air is mug d the airc ec erystal d	nswer the in, startin ail, which gh altitudes altitudes set the engin r vapor to the engin set the expansion set the expansion set the expansion the engin set the expansion the engine the enginether enginethere enginethere enginethere enginethere engineth	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v aust gases fi e. When the condense in movie state [1]. The first air [2] and irst air [1]. The first air [2] and work: work: urtner, Louis ated orange j ure's upcomin n to purchass price affect the orange crop price of FC	d generatic y]. vith [Fina condensa con	: a query to 1 Content] tion trail crence ses the t engine vapor mix troplets and the cold gases mix ter vapor when mix uld buying query mix tures eport, whice tures eport, whice tures eport, whice tures output and the buok
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for the for exmaple: [Question]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a is formed when the hot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The to albehind jets at high a set of the to exhaust gases from a between the engine exhaust and the surrou- exhaust gases to rapidly expand and cool, (Query 3]: Why does the water vapor in the arrow of the expansion of the surrou- contain a significant amount of water vapor with the cold ar at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at hig air at high altitudes, causing the water vapor with the cold er at high altitudes [3]. [Question]: In Trading Places (1983, Akro to of O at a high aptice the near he end (Answer 1]: In the movie, the main charar, Aykroyd), execute a plan to bankrupt the 1 (Answer 2]: How do Billy Ray and Louis an (Answer 2]: How do Billy Ray and Louis for Answer 3]: The Duke Brothers, who are; (Answer 2]: How does buying a lot of OJ (Answer 3]: The Duke Brothers, who are; (Paury 3]: Why does buying a lot of OJ (Answer 3]: The Duke Brothers, who are; (Final Content]: In Trading Blaces, Billy F	s complex c s complex dy know at g query), sats the query of the query is the query of the query of the query is the query of the query of the query is the query of the query	juestion n each sta in jby rei le? wmmonly are mix we e colder wmmonly are mix we with a sea and if freezes a warbulent is formeet ween the mperature of heated 1 freezes is formeet re to cond veen the mperature of the sea and places? ay Valer e re s by m he FCOJ mation al e or ange and then age.	[Question]: p of the rear point ferring the [ main and the colder: main at high a line of the colder main at high at h	"{}" and a {prime of the second secon	nswer the in, startin in, startin ail, which gh altitudes alt	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v aust gases fi e. When the condense in movie state [1]. The ing air [2], a vloses heat a movie work' artner, Louis ated orange j ure's upcomin n to purchass price after th orange crop price of FCO	d generatic y]. vith [Fina condensa condensa condensa e water v cond the j condensa	: a query to 1 Content]. 1 Co
Construct a global reasoning chain for thit the search engine based on what you alree You should generate the answer for each [ You should generate the final answer for the For example: [Question]: What causes the trail behind [Query 1]: What is the trail behind jets at [Answer 1]: The trail behind jets at high a is formed when the hot exhaust gases from to condense and freeze into tiny ice crysta (Answer 2]: The hot exhaust gases from a between the engine exhaust and the surrou- exhaust gases to rapidly expand and cool, [Query 3]: Why does the water vapor in the arrow of the twater vapor in the air cor- contain a significant amount of water vapor with the cold ar at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at high air at high altitudes, causing the water vapor with the cold er at high altitudes, it quick then freeze into ice crystals. [Final Content]: The trail behind jets at high altitudes, with the cold er at high altitudes [3]. [Query 1]: Hust it he scheme the end- [Answer 1]: In the movie, the main charar, Aykroyd), execute a plan to bankrupt the 1 Answer 2]: How do Billy Ray and Louis on [Answer 3]: The Duke Brothers, who are; Answer 3]: The Duke Brothers, who are; I han the price of FOO will remain low. H Brothers to lose a significant amount of the there of that the price of FOO will remain low. H Brothers to lose a significant amount of the remain low. H Brothers to lose a significant amount of the remover other in released [2]. The Duke Brothers, who are; Other in Content]: In Trading Places, Billy for other in gired is released [2]. The Duke I	s complex c s complex c dyery), sa he [Quesy], sa he [Quesy], sa he [Quesy], sa he [Quesy], and he cop due [he [Quesy], and he cop due [he [he [he [he [he [he [he [he [he [h	juestion n each ste infin by ref le? wmmonly are mix with se mix with se mix with se mix with se and numeric mix with this part of the second second second second ween the mperature of heated 1 freezes is formeet is formeet is formeet is formeet res by m he FCOJ mation al second se	(Question): p of the rear point ferring the [ provide the context of the context of the provide the context of the context of the rear of the context of the context of the provide the context of the context	"{}" and a soning ch Query]-{A as a contr tr air at high air is mug thitudes? i at high air is mug d the airce ce crystal d the airce d the air	nswer the in, startin nswer] pa ail, which gh altitudes altit	question, and g with [Quer irs, starting v is short for " s, causing th due to the pro- nul thinner, v aust gases fi e. When the condense in main a set enging stals [1]. The ling air [2], a vloses heat a movie work' artner, Louis tated orange j ure's upcomin n to purchass price after th orange crop price of FCO he FCOJ futt fore selling t money beca	d generatic y]. vith [Fina condensa condensa essure diff which cau: com the je hot water to liquid d exhaust e	: a query to 1 Content]. 1 Content]. 1 Content]. 1 Content]. 2 Content 2 Content

Figure 12: Prompt for generating Chain-of-Query on ELI5(in the setting without information retrieval).

Construct a global reasoning chain for this complex question [Question]." []" and answer the question, and generate a query to the search engine based on what you already know at east pot of the reasoning chain, starting with [Query]. You should generate the answer for each [Query], starting with [Answer]. You should generate the final answer for the [Question] by referring the [Query]-[Answer] pairs, starting with [Final Content]. If you don't know the answer, generate a query to the search engine based on what you already know and donot know, starting with [Unsolved Query] and please stop your generation. For example: [Question]." How many places of higher learning are in the city where the Yongle emperor greeted the person to whom the edict was addressed?" [Unsolved Query]. Who was the edict addressed to? If you don't know the answer: [Unsolved Query]. Who was the edict addressed to? If you don't know the enswer: [Answer ]: Kampa [Answer ]: [Kampa [Answer ]: [Kampa [Answer ]]: [Kampa [Answer]]: [Kampa	
[Query 2]: Where did the Yongle Emperor greet the the Karmapa? If you don't know the answer: [Unsolved Ouery]: here did the Yongle Emperor greet the the Karmapa?	
If you know the answer: [Answer 2]: Nanjing [Ouers 3]: How many nlaces of higher learning are in Nanjing ?	
If you don't hnow the answer: [Unsolved Query]: How many places of higher learning are in Nanjing? If you know the answer:	
[Answer 3]: /3 [Final Content]: The edict was addressed to Karmapa [1]. Yongle Emperor greet the Karampa in Nanjing [2]. There are 75 places of higher learning are in Nanjing [3]. So the final answer is 75.	
[Question]:"Nicholas Brody is a character on Homeland. (SUPPORTS or REFUTES)?" [Query 1]: What is Homeland? [Answer 1]: Homeland is a television series. [Query 2]: Is Nicholas Brody a character in Homeland? [Answer 2]: Yes. [Final Content]: Homeland is a television series [1]. Nicholas Brody is a character in Homeland [2]. So the final answer is SUPPORTS.	
[Question],"Brad Wilk helped co-found Rage in 1962. (SUPPORTS or REFUTES)?" [Quesy 1]: Did Brad Wilk co-found Rage? [Quesy 2]: Did Brad Wilk co-found Rage in 1962? [Answer 2]: No, Rage was founded in 1991 [Final Content]: Brad Wilk did co-found Rage [1], but not in 1962 [2]. So the final answer is REFUTES.	
[Question]:"Aristotle spent time in Athens. (SUPPORTS or REFUTES)?" [Quesry 1]: Who is Aristotle? [Answer 1]: Aristotle was a Greek philosopher. [Query 2]: Did Aristotle spend time in Athens? [Answer 2]: Yes, Aristotle studied and taught at the Academy in Athens for 20 years. [Final Content]: Aristotle was a Greek philosopher who studied and taught at the Academy in Athens for 20 years [2]. So the final answer is SUPPORTS.	
[Question]."Telemundo is a English-language television network. (SUPPORTS or REFUTES)?" [Quesy 1]: Vihat is Telemundo? [Answer 1]: Telemundo is a television network. [Quesy 2]: Is Telemundo is a Spanish-language television network? [Answer 2]: No, Telemundo is a television network [1], but it is not an English-language television network [2]. So the final answer is REFUTES.	

# Figure 13: Prompt for generating Chain-of-Query at the first round on FEVER (in the setting with information retrieval).

WWW '24, May 13-17, 2024, Singapore

1625	Construct a global reasoning chain for this complex question [Question]:"{}" and answer the question, and generate a query to the search engine based on what you already know at each step of the reasoning chain, starting with [Query].
1626	You should generate the answer for each [Query], starting with [Answer]. You should generate the final answer to judge whether to SUPPORTS or REFUTES for the [Question] by referring the [Query]-
1627	[Answer] pairs, starting with [Final Content]. For exmaple:
1628	[Question]:"SUPPORTS or REFUTES this claim?: Nicholas Brody is a character on Homeland. (SUPPORTS or REFUTES)?"
1629	[Query 1]: What is Homeland? [Answer 1]: Homeland is a television series.
1630	[Query 2]: Is Nicholas Brody a character in Homeland? [Answer 2]: Yes.
1631	[Final Content]: Homeland is a television series [1]. Nicholas Brody is a character in Homeland [2]. So the final answer is SUPPORTS.
1632	[Question]:"SUPPORTS or REFUTES this claim?: Brad Wilk helped co-found Rage in 1962. (SUPPORTS or REFUTES)?"
1633	[Query 1]. Dia Dia win Co-ronna Rage: [Answer 1]: Yes [Ouege 2]. Did Band Wills on found Bong in 10622
1634	[Answer 2]: No, Rage was founded in 1991
1635	[rinal Content]: Brad Wilk did Co-tound Rage [1], but not in 1902 [2]. So the final answer is KEPUTES.
1636	[Query 1]: Who is Aristotle? [Query 1]: Who is Aristotle?
1637	[Auswer 1]. Auswer was a Greek minisopher. [Query 2]: Did Aristotle spend time in Athens? [Auswer 2]: No. A ministrate studied and target at the Academy in Athens for 20 years.
1638	[Final Content]: Aristotle source and targin at the Academy in Athenis to 20 years. [Final Content]: Aristotle was a Greek philosopher who studied and taught at the Academy in Athens for 20 years [2]. So the final answer is SUPPORTS.
1639	
1640	[Question]:"SUPPORTS or REFUTES this claim?: Telemundo is a English-language television network. (SUPPORTS or REFUTES)?"
1641	[Query 1]: What is Telemundo? [Answer 1]: Telemundo is a television network. [Ouer: 2]: Lealonundo cos Enclich Ionemper television network?
1642	[Query 2]: IN Teleminido an English-tanguage television network? [Answer 2]: No, Telemundo is a Spanish-language television network. [Engl Constant]: Talamundo is a talavision natwork [1] but it is not an English language talavision natwork [2] So the final
1643	it mat contenty. Exchange is a television network [1], out it is not an English-language television network [2]. So the linal answer is REFUTES.
1644	

#### Figure 14: Prompt for generating Chain-of-Query on FEVER (in the setting without information retrieval).

1649	
1040	Construct a global reasoning chain for this complex [Question] : " {} "You should generate a query to the search engine based on
1649	what you already know at each step of the reasoning chain, starting with [Query].
1017	If you know the answer for [Query], generate it starting with [Answer].
1650	You can try to generate the final answer for the [Question] by referring to the [Query]-[Answer] pairs, starting with [Final
1050	Content].
1651	If you don't know the answer, generate a query to search engine based on what you already know and do not know, starting with
1051	[Unsolved Query].
1652	
1002	For exmaple:
1653	[Question]:"Is it common to see frost during some college commencements?"
1000	[Query 1]: What seasons can you expect see frost?
1654	If you don't know the answer:
	[Unsolved Query]: What seasons can you expect see frost?
1655	Instruction: Please Stop your generation.
1000	If you know the answer:
1656	[Answer I]: Winter.
1000	[Query 2]: What months do college commencements occur?
1657	II you don't know the answer:
	[onsolved Query], what months do come continencements occur?
1658	If you heave the account
	I you know the answer.
1659	[Answer 2]. Determoti, and, and sometimes June occur during winter?
	[Query 5]. Do any or December, way, and sometimes functional during winter:
1660	[Unsolved Query]: Do any of December. May, and sometimes June occur during winter?
	If you know the answer:
1661	[Answer 3]: December
	Instruction: Please Stop your generation.
1662	If you know the answer:
	[Question]:"Is it common to see frost during some college commencements?" (The answer can only be "Yes" or "No")
1663	[Final Content]: You expect see frost in Winter [1]. College commencements occur on December, May, and sometimes June [2].
	December, May, and sometimes June occur during winter [3]. So the final answer is Yes.
1664	
	[Question]:"Would a pear sink in water?"
1665	[Query 1]: What is the density of a pear?
	[Answer 1]: 0.59 g/cm^3
1666	[Query 2]: What is the density of water?
	[Answer 2]: 1 g/cm^3
1667	[Query 3]: Is 0.59 g/cm <sup>3</sup> greater than g/cm <sup>3</sup> ?
	[Answer 3]: No
1668	[Question]:"Would a pear sink in water?" (Yes or No)
	[Final Content]: The density of a pear is 0.59 g/cm <sup>3</sup> [1]. The density of water is 1 g/cm <sup>3</sup> [2]. 0.59 g/cm <sup>3</sup> is not greater than
1669	g/cm^3 [3]. So the final answer is No.
	L
1670	

#### Figure 15: Prompt for generating Chain-of-Query at the first round on StragegyQA (in the setting with information re-trieval).

Construct a global reasoning chain for this complex [Question] : " {} " You should generate a query to the search engine based on			
what you already know at each step of the reasoning chain, starting with [Query].			
If you know the answer for [Query], generate it starting with [Answer].			
You can try to generate the final answer for the [Question] by referring to the [Query]-[Answer] pairs, starting with [Final			
Content].			
If you don't know the answer, generate a query to search engine based on what you already know and do not know, starting with			
[Unsolved Query].			
For example:			
[Question]: "Where do greyhound buses that are in the birthplace of Spirit If's performer leave from? "			
[Query 1]: Who is the performer of Spirit If ?			
If you don't know the answer:			
[Unsolved Query]: Who is the performer of Spirit If ?			
If you know the answer:			
[Answer 1]: The performer of Spirit If is Kevin Drew.			
[Query 2]: Where was Kevin Drew born?			
If you don't know the answer:			
[Unsolved Query]: Where was Kevin Drew born?			
If you know the answer:			
[Answer 2]: Toronto.			
[Query 3]: Where do greyhound buses in Toronto leave from?			
If you don't know the answer:			
[Unsolved Query]: Where do greyhound buses in Toronto leave from?			
If you know the answer:			
[Answer 3]: Toronto Coach Terminal.			
[Final Content]: The performer of Spirit If is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in			
Toronto leave from Toronto			
Coach Terminal [3]. So the final answer is Toronto Coach Terminal.			
[Question]:"Which magazine was started first Arthur's Magazine or First for Women?"			
[Query 1]: When was Arthur's Magazine started?			
[Answer 1]: 1844.			
[Query 2]: When was First for Women started?			
[Answer 2]: 1989			
[Final Content]: Arthur's Magazine started in 1844 [1]. First for Women started in 1989 [2]. So Arthur's Magazine was started			
first. So the answer is Arthur's Magazine.			

# Figure 16: Prompt for generating Chain-of-Query at the first round on HotpotQA, Musique, WikiMultiHopQA, zsRE and T-REx (in the setting with information retrieval).

Construct a global reasoning chain for this complex question [Question]." {}" and answer the question, and generate a query to the search engine based on what you already know at each step of the reasoning chain, starting with [Query]. You should generate the answer for each [Query], starting with [Answer]. You should generate the final answer to judge whether to SUPPORTS or REFUTES for the [Question] by referring the [Query]. [Answer] pairs, starting with [Final Content]. For exmaple:
[Question]."SUPPORTS or REFUTES this claim?: Nicholas Brody is a character on Homeland. (SUPPORTS or REFUTES)?" [Query 1]: What is Homeland? [Answer 1]: Homeland is a television series.
[Query 2]: Is Nicholas Brody a character in Homeland?
[Answer 2]: Tes. [Final Content]: Isomeland is a television series [1]. Nicholas Brody is a character in Homeland [2]. So the final answer is SUPPORTS.
[Question]: SUPPORTS or REFUTES this claim?: Brad Wilk helped co-found Rage in 1962. (SUPPORTS or REFUTES)?" [Query 1]: Did Brad Wilk co-found Rage? [Answer 1]: Yes
[Query 2]: Did Brad Wilk co-found Rage in 1962?
[Answer 2]: No, Rage was founded in 1991
[Final Content]: Brad Wilk did co-found Rage [1], but not in 1962 [2]. So the final answer is REFUTES.
[Question]:"SUPPORTS or REFUTES this claim?: Aristotle spent time in Athens. (SUPPORTS or REFUTES)?" [Query 1]: Who is Aristotle?
[Answer 1]: Aristotle was a Greek philosopher. [Ouery 2]: Did Aristotle spend time in Athens?
[Answer 2]: Yes, Aristotle studied and taught at the Academy in Athens for 20 years.
[Final Content]: Aristotle was a Greek philosopher who studied and taught at the Academy in Athens for 20 years [2]. So the final answer is SUPPORTS.
[Question]:"SUPPORTS or REFUTES this claim?: Telemundo is a English-language television network. (SUPPORTS or REFUTES)?"
[Query 1]: What is Telemundo?
[Answer 1]: Telemundo is a television network. [Ouery 2]: Is Telemundo an English-language television network?
[Answer 2]: No, Telemundo is a Spanish-language television network.
[Final Content]: Telemundo is a television network [1], but it is not an English-language television network [2]. So the final answer is REFUTES.

# Figure 17: Prompt for generating Chain-of-Query on FEVE (in the setting without information retrieval).

		1	
-			
		L	
-	F	2	
3	F	 	
3	F	 	
Ξ	F	 	
	F	 ₹	
	I	 ₹	
3	F	 ₹	
E	F		
E	F	 1	
3	F	 R	
2	I		
E	I	2	
	ŀ	 R	
E	F		

### WWW '24, May 13-17, 2024, Singapore

1741	Construct a global reasoning chain for this complex question [Question]:"{}" and answer the question, and generate a query to the	
	search engine based on what you already know at each step of the reasoning chain, starting with [Query].	
1742	You should generate the answer for each [Query], starting with [Answer]. You should generate the final answer for the [Question] by referring the [Query]-[Answer] pairs, starting with [Final Content].	
1743	For avmanle1:	
	[Question]: "Do hamsters provide food for any animals?"	
1744	[Query 1]: What types of animal are hamsters?	
1745	[Answer 1]: Hamsters are prey animals.	
1/45	[Query 2]: Do prey provide food for any other animals?	
1746	[Answer 2]: Yes	
	[Question]: Do namisters provide food for any animals? (res or two) [Final Content]: Hamsters are new animals [1]. Pray reprovide food for other animals [2]. So the final answer is Vas	
1747	[1 mai concent]. Hanseets are prey annuas [1]. I tey provide root for outer annuais [2]. 50 the final answer is res.	
	For exmaple2:	
1748	[Question]:"Could Brooke Shields succeed at University of Pennsylvania?"	
1740	[Query 1]: What college did Brooke Shields go to?	
1749	[Answer 1]: Princeton University	
1750	[Query 2]: Out of all coneges in the US, now is Princeton University ranked	
1,00	[Auswei 2]. Frinceton is fainced as the functional contege to 5 news. [Outery 3]: Is the ranking of University of Pennsylvania similar to Princeton University?	
1751	[Answer3]: Yes	
	[Question]:"Could Brooke Shields succeed at University of Pennsylvania?" (Yes or No)	
1752	[Final Content]: Brooke Shields went to Princeton University [1]. Princeton is ranked as the number 1 national college by US	
1750	news [2]. The ranking of University of Pennsylvania similar to Princeton University [3]. So the final answer is Yes.	
1/53	For average [2]	
1754	FOI exhiptes. [Ouestion]:"Is it common to see frost during some college commencements?"	
	[Ouery 1]: What seasons can you expect see frost?	
1755	[Answer 1]: Winter.	
	[Query 2]: What months do college commencements occur?	
1756	[Answer 2]: December, May, and sometimes June.	
1757	[Query 3]: Do any of December, May, and sometimes June occur during winter? [Anowy 3]: No.	
1/5/	[Auswer 5]. 165. [Ouestion]:"Is it common to see frost during some college commencements?" (Yes or No)	
1758	[Final Content]: You expect see frost in Winter [1]. College commencements occur on December. May, and sometimes June [2].	
1,00	December, May, and sometimes June occur during winter [3]. So the final answer is Yes.	
1759		
	For exmaple4:	
1760	[Question]: "Would a pear sink in water?"	
1771	[Query 1]. what is the density of a peak: [Answer 1]: 0.52 o/cm^3	
1/01	[Query 2]: What is the density of water?	
1762	[Answer 2]: 1 g/cm^3	
1702	[Query 3]: Is 0.59 g/cm <sup>3</sup> greater than g/cm <sup>3</sup> ?	
1763	[Answer 3]: No	
	[Question]: "Would a pear sink in Water?" (Yes or No) [Evial Contrast]. The density of a pear is 0.50 g/mm21[1]. The density of water is 1 g/mm21[2] 0.50 g/mm23 is not apparent than	
1764	[Final Content]. The density of a pear is 0.59 gent 5 [1]. The density of water is 1 gent 5 [2], 0.59 gent 5 is not greater than a/cm/3 [3]. So the final answer is No.	
17/5		
1/05	For exmaple5:	
1766	[Question]:"Could a llama birth twice during War in Vietnam (1945-46)?"	
1,00	[Query 1]: How long did the Vietnam war last?	
1767	[Answer 1]: Around 6 months.	
	[Query 2]. How long is name gestational period: [Answer 2]: The sectation period for a llama is 11 months	
1768	[Ouery 3]: What is 2 times 11 months?	
17/0	[Answer 3]: 22 months.	
1/69	[Query 4]: Is 6 months longer than 22 months?	
1770	[Answer 4]: No	
1110	[Question]: "Could a llama birth twice during War in Vietnam (1942–46)?" (Yes or No) Existe Contrastly Vietnam was had respired 6 growther ll. The organized for a llama in 11 months [2], 2 times 11 months in 22	
1771	[Phat Content]. Vientian wai tast around o montus [1]. The gestation period for a name is 11 montus [2]. 2 times 11 montus is 22 months [3] 6 months is not honer than 22 months [4]. So the final answer is No.	
	norma (5), o montas a not tonget ann 22 montas (1), so are man ansite ta sto.	
1772	For exmaple6:	
1772	[Question]:"Hydrogen's atomic number squared exceeds number of Spice Girls?"	
1//3	[Query 1]: What is the atomic number of hydrogen?	
1774	[Answer 1]: Hydrogen is the first element and has an atomic number of one. [Ouerv 2]: How many people are in the Spice Girls hand?	
	[Answer 2]: The Snice Girk has five members	
1775	[Query 3]: Is the square of 1 greater than 5?	
	[Answer 3]: No	
1776	[Question]:"Hydrogen's atomic number squared exceeds number of Spice Girls?" (Yes or No)	
1777	[Final Content]: Hydrogen is the first element and has an atomic number of one [1]. The Spice Girls has five members [2]. The	
1///	square of 1 is not greater than 5131. So the linal answer is No.	1

Construct a global reasoning chain for this complex question [Ouestion]:"{}" and answer the question, and generate a query to the

Figure 18: Prompt for generating Chain-of-Query on StrategyQA (in the setting without information retrieval).

#### Example of the Interaction between IR and A.5 LLM

We use Figure 20 and Figure 21 as an example to show the details of the interaction between IR and LLM. In this example, IR and LLM perform three rounds of interaction. IR verifies and corrects the answers to the first two sub-questions and provides the answer to the last sub-question. In each round, LLM re-generate the new CoQ according to the feedback of IR.

# [Query 2]: Why do the hot exhaust gases mix with the colder air at high altitudes? If you don't know the answ [Unsolved Ouerv]: Why do the hot exhaust gases mix with the colder air at high altitudes? If you know the If you know the answer: If you know the disevert is the explosion of the end of the e If you don't know the answer If yous know the answer: If you know the answer: [Answer 3]: The water vapor in the air condenses and freezes into ice crystals? If you know the maxwer: [Answer 3]: The water vapor in the air condenses and freezes into ice crystals because the exhaust gases from the jet engine contain a significant amount of water vapor, which is heated to high temperatures in the engine. When the hot water vapor mixes with the cold air an high altitudes, i quickly loses heat and energy, causing the water vapor to condense into liquid droplets and them freeze into ice crystals. [Final Content]: The trail behind jets at high altitude is formed when the hot exhaust gases from a jet engine mix with the colder air at high altitudes, causing the water vapor in the air to condense and freeze into univ ice crystals. The exhaust gases mix with the colder air due to the pressure difference between the engine exhaust and the surrounding air, and the water vapor condenses and freezes heates it is heated to high temperatures in the engine and then quickly loses heat and energy when mixed with the cold air at high altitudes. [Unsolved Query]: Why does the water vapor in the air condense and freeze into ice crystals? If you know the answer: [Question]: In Trading Places (1983, Akroyd/Murphy) how does the scheme at the end of the movie work: why would we are used lot of OJ at a high price ruin the Dake Brothers? [Query 1]: What is the scheme at the end of Trading Places? [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Murphy), and his partner, Louis Winthorpe III (Dan [Answer 1]: In the movie, the main character, Billy Ray Valentine (Eddie Mur ion]: In Trading Places (1983, Akroyd/Murphy) how does the scheme at the end of the movie work? Why would buying a

onstruct a global reasoning chain for this complex [Question]: "{} "You should generate a query to the search engine based a what you already know at each step of the reasoning chain, starting with [Query]. [you know the asswer for [Query], generate it starting with [Answer] ou can try to generate the final answer for the [Question] by referring to the [Query]-[Answer] pairs, starting with [Final

Content]. If you don't know the answer, generate a query to search engine based on what you already know and do not know, starting with [Unsolved Query].

[Answer 1]: The trail behind jets at high altitude is commonly referred to as a contrail, which is short for "condensation trail." It

is formed when the hot exhaust gases from a jet engine mix with the colder air at high altitudes, causing the water vapor in the ai

For exmaple: [Question]:"What causes the trail behind jets at high altitude? [Query 1]: What is the trail behind jets at high altitude?

If you know the answer

to condense and freeze into tiny ice crystals

If you don't know the answer: [Unsolved Query]: What is the trail behind jets at high altitude?

market [Query 2]: How do Billy Ray and Louis manipulate the FCOJ futures market? (2000) = 21: Billy Ray and Louis obtain insider information about the Department of Agriculture's upcoming crop report, which indicates that a harsh winter has destroyed much of the orange crop. They use this information to purchase FCOI futures contracts at a low price before the report is released, and then they sell the contracts at a high price after the report's release. connects in a low price costs, the epoints received, and one may sent an comments in a migraprice and the report's recease, when the market that sersponded to the news of the crop damage. [Query 3]: Why does buying a lot of OI at a high price ruin the Duke Brothers? [Answer 3]: The Duke Brothers, who are also investing in the FCOD market, have bet that the orange crop will be abundant and [Answer 3]: The Duke Brothers, who are also investing in the FCOJ market, have bet that the orange crop will be abundant and that the price of FCOJ will remain low. However, Billy Ray and Louis's scheme drives up the price of FCOJ, causing the Duke Brothers to lose a significant amount of money and ultimately leading to their downfall. [Final Content]: In Trading Places, Billy Ray Valentine and Louis Winthorpe III manipulate the FCOJ futures market by obtaining insider information about the crop report and purchasing contracts at a low price before selling them at a higher price after the report released. The Duke Brothers, who have also invested in the market, Lose more because they bet on an abundant orange crop and low FCOJ prices. However, Billy Ray and Louis's scheme causes the price of FCOJ to rise, which mins the Duke Brothers and leads to their downfall. abundant orange crop and low FCOJ prices. Howev ruins the Duke Brothers and leads to their downfall

# Figure 19: Prompt for generating Chain-of-Query at the first round on ELI5 (in the setting with information retrieval).

Anon.

WWW '24, May 13-17, 2024, Singapore

Question: Where do greyhound buses that are in the birthplace of Spirit If...'s performer leave from?

1857	Question: Where do greyhound buses that are in the birthplace of Spirit If's performer leave from?				
	Round 1 LLM CoQ 1 generation				
1858	[Query 1]: "Who is the performer of Spirit If?"				
1859	[Answer 1]: "Casey Morton Cott"				
	[Query 2]:"What is the place of birth of Casey Morton Cott?				
1860	[Answer 2]: Cuyanoga County [Linsolved Query]: "Where do greyhound buses leave from in Cuyahoga County?"				
1861					
1001	Round 1 IR Verification				
1862	For Node 1 of CoQ 1:				
1863	[Query 1]: "Who is the performer of Spirit If?"				
1005	[Answer 1]: "Casey Morton Cott"				
1864					
1865	Secret ("On-Whe is the performent of Spirit (6, 2")				
1005	Top-1 Passage				
1866	Spirit If is the debut colo Reader If ("Kevin Drew" not match "Casey				
1867	album by Kevin Drew. It Model → Answer is : Kevin Drew Morton Cott" && f > threshold) ?				
1007	was released on September				
1868	18, 2007 Yes, construct a prompt to feed back to LLM				
1869	ļ				
1007	According to this Reference, the answer for "Who is the performer of Spirit If?" should be "Kevin Drew",				
1870	you can change your answer based on the Reference and continue constructing the reasoning chain to give				
1871	the final answer for [Question]: "Where do greynound buses that are in the birthplace of Spirit it s				
10/1	Reference: Spirit If is the debut solo album by Kevin Drew. It was released on September 18, 2007				
1872					
1873	Round 2 LLM CoQ 2 generation				
1075	[Query 1]: "Who is the performer of Spirit If?"				
1874	[Answer 1]: "Kevin Drew"				
1875	[Answer 2]: "Vancouver"				
1075	[Unsolved Query]:"Where do greyhound buses leave from in Vancouver?"				
1876					
1877	Round 2 IR Verification				
	[Ouery 1]: "Who is the performer of Spirit If?"				
1878	[Answer 1]: "Kevin Drew"				
1879	If Q="Who is the performer of Spirit If?" has been processed ? Yes				
10//	For Node 2 of CoQ 2:				
1880	[Query 2]: "What is the place of birth of Kevin Drew?"				
1881	If Q="What is the place of birth of Kevin Drew?" has been processed?				
	No				
1882	Search ("O=What is the place of high of Keyin Drow?"")				
1883	Search ( Q-what is the place of birth of Kevin Drewr )				
	Top-1 Passage				
1884	Kevin Drew (born Model Answer is : Toronto If ("Toronto" not match				
1885	September 9, 1976 "Vancouver" && f > threshold) ?				
1886	Yes, construct a prompt to feed back to LLM				
1887					
1007	According to this Reference, the answer for "What is the place of birth of Kevin Drew?" should be "Toronto",				
1888	you can change your answer based on the Reference and continue constructing the reasoning chain to give				
1889	performer leave from?"				
	Reference: Kevin Drew (born September 9, 1976 in Toronto)				
1890	[				
1891					

Round 3 LLM CoQ 3 generation [Query 2]:"What is the place of birth of Kevin Drew?" [Answer 2]: "Toronto" [Unsolved Query]:"Where do greyhound buses leave from in Toronto?"
Densel 2 ID (Densels to the
For Node 1 of CoD 2:
[Query 2]:"What is the place of hirth of Kevin Drew?
[Answer 2]: "Toronto"
If Q="Mithat is the place of binth of You're Draw 2" has been proceeded 2 Ver
For Node 2 of CoD 3:
[Unsolved Query]:"Where do greyhound buses leave from in Toronto?"
If Q="Where do greyhound buses leave from in Toronto?" has been processed?
No
Search ("O=Where do greyhound buses leave from in Toronto?")
Top-1 Passage
The Toronto Coach Terminal is the central bus Reader
station for inter-city services in Toronto, Ontario, Model Answer is : Toronto Coach Terminal
Canada when it was leased out in its entirety to
bus lines Coach Canada and Greyhound Canada
construct a prompt to feed back to LLM
"Toronto Coach Terminal", you can give your answer based on the Reference and continue constructing the reasoning chain to give the final answer for [Question]: "Where do greyhound buses that are in the birthplace of Spirit If's performer leave from?" Reference: The Toronto Coach Terminal is the central bus station for inter-city services in Toronto, Ontario, Canada when it was leased out in its entirety to bus lines Coach Canada and Greyhound Canada
Final Basside Tracing
Primar Result: Trachig
You can try to generate the final answer for the [Question] by referring to the [Query]-[Answer] pairs, starting
with [Final Content].
[Query 1]: "Who is the performer of Spirit If?"
[Answer 1]: "Kevin Drew"
[Query 2]:"What is the place of birth of Kevin Drew?
[Answer 2]: "Toronto"
10. Jon 21. When a de graute and bused leave trans in Terente 3/
[Query 3]: "Where do greyhound buses leave from in Toronto?"
[Query 3]: "Where do greyhound buses leave from in Toronto?" [Answer 3]: "Toronto Coach Terminal "
[Query 3]:"Where do greyhound buses leave from in Toronto?" [Answer 3]: "Toronto Coach Terminal " Output:
[Query 3]:"Where do greyhound buses leave trom in Toronto?" [Answer 3]: "Toronto Coach Terminal " Output: The performer of Spirit If is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in
[Query 3]:"Where do greyhound buses leave trom in Toronto?" [Answer 3]: "Toronto Coach Terminal " Output: The performer of Spirit If is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in Toronto leave from Toronto Coach Terminal [3]. So the final answer is Toronto Coach Terminal.
[Query 3]:"Where do greyhound buses leave from in Toronto?" [Answer 3]: "Toronto Coach Terminal " Output: The performer of Spirit If is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in Toronto leave from Toronto Coach Terminal [3]. So the final answer is Toronto Coach Terminal.
[Query 3]: "Where do greyhound buses leave trom in Toronto?"         [Answer 3]: "Toronto Coach Terminal "         Output:         The performer of Spirit If is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in Toronto leave from Toronto Coach Terminal [3]. So the final answer is Toronto Coach Terminal.         [1] spirit If is the debut solo album by Kevin Drew. It was released on September 18, 2007
[Query 3]:"Where do greyhound buses leave trom in Toronto?" [Answer 3]: "Toronto Coach Terminal " Output: The performer of Spirit f is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in Toronto leave from Toronto Coach Terminal [3]. So the final answer is Toronto Coach Terminal. [1] Spirit f is the debut solo album by Kevin Drew. It was released on September 18, 2007 [2] Kevin Drew (born September 9, 1976 in Toronto) [3] Toronto Coach Terminal I. the caretal bus station for inter-city services in Toronto Dotatio Coach
<ul> <li>[Query 3]:"Where do greyhound buses leave trom in Toronto?"</li> <li>[Answer 3]: "Toronto Coach Terminal "</li> <li>Output:</li> <li>The performer of Spirit If is Kevin Drew [1]. Kevin Drew was born in Toronto [2]. Greyhound buses in Toronto leave from Toronto Coach Terminal [3]. So the final answer is Toronto Coach Terminal.</li> <li>[1] Spirit If is the debut solo album by Kevin Drew. It was released on September 18, 2007</li> <li>[2] Kevin Drew (born September 9, 1976 in Toronto)</li> <li>[3] The Toronto Coach Terminal is the central bus station for inter-city services in Toronto, Ontario, Canada</li> </ul>

Figure 21: Details (Round 3, Final).

Figure 20: Details (Round 1, 2).