

---

000 PLANNING IN STRAWBERRY FIELDS: EVALUATING  
001 AND IMPROVING THE PLANNING AND SCHEDULING  
002 CAPABILITIES OF LRM o1  
003  
004  
005

006 **Anonymous authors**

007 Paper under double-blind review  
008  
009

010  
011 ABSTRACT  
012

013 The ability to plan a course of action that achieves a desired state of affairs  
014 has long been considered a core competence of intelligent agents and has been  
015 an integral part of AI research since its inception. With the advent of large  
016 language models (LLMs), there has been considerable interest in the question  
017 of whether or not they possess such planning abilities, but—despite the slew of  
018 new private and open source LLMs since GPT3—progress has remained slow.  
019 OpenAI claims that their recent o1 (Strawberry) model has been specifically  
020 constructed and trained to escape the normal limitations of autoregressive  
021 LLMs—making it a new kind of model: a Large Reasoning Model (LRM). In  
022 this paper, we evaluate the planning capabilities of two LRMs (o1-preview  
023 and o1-mini) on both planning and scheduling benchmarks. We see that  
024 while o1 does seem to offer significant improvements over autoregressive  
025 LLMs, this comes at a steep inference cost, while still failing to provide  
026 any guarantees over what it generates. We also show that combining o1  
027 models with external verifiers—in a so-called LRM-Modulo system—guarantees  
028 the correctness of the combined system’s output while further improving  
029 performance.

030  
031 1 INTRODUCTION  
032

033 The recent release of OpenAI’s o1 (Strawberry) (OpenAI, 2024a) brings with it the op-  
034 portunity to freshly evaluate the progress of large pre-trained AI models on planning and  
035 scheduling benchmarks. Unlike the Large Language Models (LLMs) which came before  
036 it—which can roughly be viewed as approximate retrievers—o1 seems to have been trained to be  
037 an approximate reasoner, capable of scaling the amount of compute it uses depending on the  
038 query.<sup>1</sup> Following OpenAI<sup>2</sup>, we draw a distinction between previous Large Language Models  
039 (LLMs) and o1, a Large Reasoning Model (or LRM), as its new (unknown) architecture,  
040 operation, and capabilities all seem to be fundamentally different from those of vanilla LLMs,  
041 likely during both the pre-training phase and at inference time. Our aim in this paper is  
042 twofold: to comprehensively evaluate the performance of o1 on established planning and  
043 scheduling benchmarks as well as more difficult extensions; and to demonstrate how to  
044 provide correctness guarantees and boost performance by embedding an LRM in a loop  
045 with a sound verifier, in a similar vein to the LLM-Modulo framework (Kambhampati et al.,  
046 2024).

---

047 <sup>1</sup>We speculate that the complete system learns to improve its ability to make appropriate Chain-  
048 of-Thought (CoT) moves useful for reasoning in a pretraining RL step with synthetic data, and does  
049 inference time prompt-specific rollouts; see Appendix B. In other words, it may be an RL-trained  
050 system in the same vein as AlphaGo, but where the ‘moves’ being generated and evaluated are  
051 Chains of Thought. Note that this is very different use of RL from that in RLHF, which can be  
052 seen as a fine tuning stage for an LLM that keeps the inference step unchanged.

053 <sup>2</sup>Per the blogpost announcing the model: "A new series of reasoning models [...] for complex  
reasoning tasks this is a significant advancement and represents a new level of AI capability. Given  
this, we are resetting the counter back to 1 and naming this series OpenAI o1."(OpenAI, 2024b)

---

054 To properly evaluate this new kind of model and understand its abilities and limitations will  
055 require new tools and evaluation methods, especially if details of the overall model structure  
056 are kept secret and internal traces remain inaccessible to outside researchers.<sup>3</sup> In this paper,  
057 we evaluate performance on established benchmarks and compare to previous state-of-the-art  
058 results, extending these benchmarks to more difficult problems when possible and necessary.  
059 For planning, we use PlanBench (Valmeekam et al., 2023), which consists of both specific  
060 test sets and a suite of tools intended for evaluating language models on arbitrary IPC  
061 planning domains. To evaluate scheduling capabilities, we test on TravelPlanner (Xie et al.,  
062 2024), on the three domains from Natural Plan (Zheng et al., 2024), and on graph coloring  
063 problems (Stechly et al., 2024b).

064 Using these benchmarks as our basis for analysis, we investigate the performance jump  
065 that LRMs from the o1 family promise. We then demonstrate how these benchmarks can  
066 be elaborated on in order to remain relevant metrics for LRMs. We argue that, to be  
067 complete, new approaches to measuring LRM reasoning capabilities must take into account  
068 efficiency, cost, and guarantees. We also note the steep inference cost of LRMs and discuss  
069 the tradeoffs between using LLMs vs LRMs, arguing that in some cases an LLM-Modulo  
070 (Kambhampati et al., 2024) approach may be significantly cheaper than o1 models for  
071 comparative performance, and with guarantees. Our results also show that that same  
072 LLM-Modulo approach can indeed be adapted to LRMs to further improve their performance  
073 and provide guarantees. In essence, LRMs can replace LLMs as significantly better—but still  
074 fallible—generators in the LLM-Modulo framework.

## 075 2 BACKGROUND AND RELATED WORK

076  
077  
078 Though they are trained as text completion systems, Large Language Models (LLMs) have  
079 shown some promise on many other tasks. Initial claims were wildly positive, claiming they  
080 are general purpose reasoning systems (Bubeck et al., 2023), especially when prompted in  
081 just the right way (Kojima et al., 2022; Wei et al., 2022), but later results showed that their  
082 seeming reasoning capabilities are brittle and break down even in simple domains (Luo et al.,  
083 2023; Dziri et al., 2024; Stechly et al., 2024b; Qian et al., 2022) and may be attributable to  
084 dataset contamination (Roberts et al., 2023). In planning, Valmeekam et al. (2023) showed  
085 that LLMs fail even on problems as trivial as three block stacking.

086 Based on what little has been revealed by OpenAI, o1 seems to be a new class of model (a  
087 Large Reasoning Model or LRM), designed to combine the fuzzy language capabilities of  
088 LLMs with some implementation of approximate reasoning. With this in mind, we believe it  
089 is time to bring up the same questions that were asked about LLMs for these LRMs. We use  
090 benchmarks from the LLM literature, extending them where possible and necessary to show  
091 how well and how robustly o1 does or doesn't perform on various planning and scheduling  
092 tasks.

### 093 2.1 DOMAINS: PLANNING

094  
095 The LLM literature abounds with claims of the 'emergent' planning capabilities of  
096 LLMs (Huang et al., 2022a). However, closer inspection reveals that many of the em-  
097 pirical results supporting these claims comes from evaluations on simpler, commonsense  
098 domains, such as ALFworld (Yao et al., 2023), BEHAVIOR (Srivastava et al., 2022), kitchen  
099 environments (Ahn et al., 2022; Huang et al., 2022b), and virtual home (Huang et al., 2022a).  
100 Not only do the the instances tested on tend to have fewer interactions, but many of them  
101 conflate reactive acting and deliberative planning (Ghallab et al., 2016).

102 In contrast, we focus on classical planning problems, or STRIPS planning problems, which  
103 are a formalism for automated planning in discrete, deterministic spaces. To define a  
104 planning problem, we specify an *initial state*, a *domain*, and a *goal*. The domain contains all  
105 relevant information about the types of objects that may exist and the allowable actions  
106 from any given state, specified by defining the preconditions and effects of each named

---

107 <sup>3</sup>There are reports that OpenAI is threatening to revoke access to o1 from anyone who tries to  
extract internal reasoning traces (Edwards, 2024).

Domain	Shots	Claude Models		OpenAI GPT-4 Models				LLaMA Models		Gemini Models	
		Claude 3.5 (Sonnet)	Claude 3 (Opus)	GPT-4o	GPT-4o-mini	GPT-4	GPT-4 Turbo	LLaMA 3.1 405B	LLaMA 3 70B	Gemini 1.5 Pro	Gemini 1 Pro
Blocksworld	One Shot	<b>346/600</b> (57.6%)	289/600 (48.1%)	170/600 (28.3%)	49/600 (8.1%)	206/600 (34.3%)	138/600 (23%)	284/600 (47.3%)	76/600 (12.6%)	101/600 (16.8%)	68/600 (11.3%)
	Zero Shot	329/600 (54.8%)	356/600 (59.3%)	213/600 (35.5%)	53/600 (8.8%)	210/600 (34.6%)	241/600 (40.1%)	<b>376/600</b> (62.6%)	205/600 (34.16%)	143/600 (23.8%)	3/600 (0.5%)
Mystery Blocksworld	One Shot	19/600 (3.1%)	8/600 (1.3%)	5/600 (0.83%)	0/600 (0%)	<b>26/600</b> (4.3%)	5/600 (0.83%)	21/600 (3.5%)	15/600 (2.5%)	-	2/500 (0.4%)
	Zero Shot	0/600 (0%)	0/600 (0%)	0/600 (0%)	0/600 (0%)	1/600 (0.16%)	1/600 (0.16%)	<b>5/600</b> (0.8%)	0/600 (0%)	-	0/500 (0%)

Table 1: Performance on 600 instances from the Blocksworld and Mystery Blocksworld domains across large language models from different families, using both zero-shot and one-shot prompts. Best-in-class accuracies are bolded.

action. Problems and domains are represented in the flexible PDDL (Planning Domain and Definition Language) framework (McDermott et al., 1998). Solutions to PDDL problems are correct plans—sequences of actions executable from the initial state which arrive at a goal-satisfying final state. These are problems in which the the planner already knows all relevant facts about the world and which actions are possible—only deliberation is required.

In the current work, we extend the STRIPS planning evaluation pipeline first introduced by Valmeekam et al. (2024). This benchmark provides an extensible suite of domains and tools for evaluating general models, a collection of static test sets across these domains, and ways of easily generating additional instances of problems in these domains. We draw on these static test sets to provide our initial o1 planning results on Blocksworld—a commonsense domain from the International Planning Competition (IPC, 1998)—and Mystery Blocksworld—an obfuscated version of the same. We also extend these sets to harder instances and examine performance on problems from both Logistics and Sokoban, two other well-known IPC domains.

o1’s performance on PlanBench is especially interesting in light of the fact that the benchmark remains very challenging for vanilla LLMs (massive transformer models which have been fine-tuned via RLHF). The lackluster performance of LLMs on even the easiest static test set leads us to continue to believe that planning cannot be generally and robustly solved by approximate retrieval alone. In Table 1, we present the results of running current and previous generation LLMs on a static test set of 600 three to five block Blocksworld problems, as well as on a set of 600 semantically identical but syntactically obfuscated instances which Valmeekam et al. (2024) calls Mystery Blocksworld.

Across the models tested, the best performance on regular Blocksworld is achieved by LLaMA 3.1 405B with 62.6% accuracy. Despite the underlying problems being identical, Mystery Blocksworld performance lags far behind—no LLM achieves even 5% on our test set—and performance on one version of the domain does not clearly predict performance on the other. We do not provide Mystery Blocksworld data for Gemini 1.5 Pro only because we haven’t been able to generate it. The model refuses to produce any output, instead claiming that responding to these queries would be harmful. We include this output in Appendix D. Vanilla language models perform better when tested on natural language versions of prompts rather than PDDL (Valmeekam et al., 2024), even though natural language can introduce uncertainty due to polysemanticity and syntactic ambiguity. To make our comparisons "fair" for the models being tested, the results we report are the higher percent accuracy natural language prompting numbers, and can be considered an upper bound on LLM performance on PDDL-specified problems.

PlanBench does not explicitly take efficiency into consideration. As the time taken by a vanilla LLM to produce some output is only dependent on the length of that output, but otherwise independent of the semantic content or difficulty of the instance, this may not have particularly mattered in evaluations at the time. However, as LRMs adaptively vary their time taken and dollar cost per instance in response to the input, measuring efficiency

Table 2: Performance and average time taken on 600 instances from the Blocksworld, Mystery Blocksworld and Randomized Mystery Blocksworld domains and 200 instances from the Logistics and Randomized Logistics domains by OpenAI’s o1 family of large reasoning models and Fast Downward (F.D.)

Total Instances	Domain	Shots	Instances correct			Average Time Taken (in secs)		
			o1-preview	o1-mini	F.D (satisficing)	o1-preview	o1-mini	F.D (satisficing)
600	Blocksworld	Zero Shot	97.8%	56.6%	100%	40.43	10.84	0.12
600	Mystery Blocksworld	One Shot	41.6%	-	100%	82.03	-	0.12
		Zero Shot	52.8%	19.1%	100%	83.37	35.54	0.12
600	Randomized Mystery Blocksworld	Zero Shot	37.3%	3.5%	100%	111.11	55.40	0.12
200	Logistics	Zero Shot	94%	-	100%	84.07	-	0.13
200	Randomized Mystery Logistics	Zero Shot	52%	-	100%	167.41	-	0.13

has become much more important. As a comparison point between LRMs and LLMs, we compute prompting costs across models and present them in Table 4.

We also find that, contrary to previous claims, one-shot prompting is *not* a strict improvement over zero-shot. In fact, for many models it seems to do significantly worse!<sup>4</sup> This is most notable in our tests of LLaMA family models.

We relegate additional discussion of the particulars of LLM performance to appendix A.

## 2.2 DOMAINS: SCHEDULING

More recent text-based benchmarks have provided full, static descriptions of their domains, initial states, and goals. However, many of these, despite the word "planning" in their titles, would be better characterized as testing scheduling abilities (Ghallab et al., 2016). Classical planning problems are canonical graph search problems which are PSPACE-complete. Scheduling problems are only NP-Hard (Carlier, 1982), and mainly revolve around resource allocation. These problems are equivalent to constraint satisfaction problems, and thus easier than the planning problems we describe above.

We evaluate o1 on three scheduling benchmarks on which LLMs have failed. Zheng et al. (2024)’s Natural Plan benchmark consists of three scheduling domains: trip planning, calendar scheduling, and meeting planning. Xie et al. (2024)’s Travel Planning benchmark consists of a large dataset of travel information (flights, accommodations, restaurants, etc.) with prompts that ask the model to create a three to seven day itinerary based on natural language instructions. Finally, Stechly et al. (2024b) translate graph coloring, a classical constraint satisfaction problem, into natural language prompts and evaluates GPT-4’s accuracy on these problems. We take their test set and extend it to more difficult instances.

## 3 FROM APPROXIMATE RETRIEVAL TO APPROXIMATE REASONING

Many researchers have argued that "standard" autoregressive LLMs generate outputs via approximate retrieval, and that, while they show impressive performance on a range of System 1 tasks, they are unlikely to achieve the more System 2-like approximate reasoning capabilities that are critical for planning tasks (c.f. (Kambhampati, 2024)). From our

<sup>4</sup>While the reverse is generally true for Mystery Blocksworld problems, it’s important to note that the performance of vanilla LLMs on Mystery Blocksworld has consistently and uniformly been poor (the same as it was when this benchmark was first released), so those results do not provide too clear a picture. Most models do not solve even a single instance in zero-shot mode, and only one (LLaMA 3.1 405B) manages more than one.

216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
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

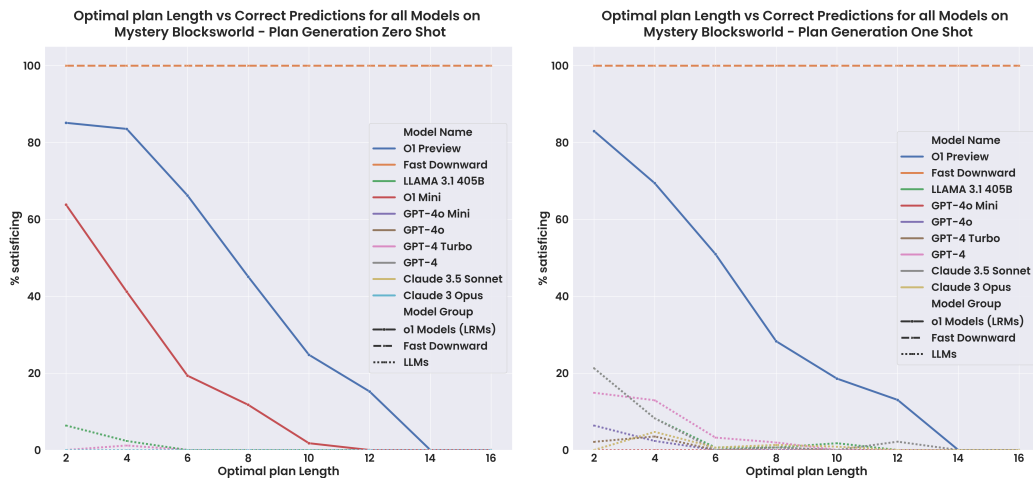


Figure 1: These examples are on Mystery Blocksworld. Fast Downward, a domain-independent planner (Helmert, 2006) solves all given instances near-instantly with guaranteed perfect accuracy. LLMs struggle on even the smallest instances. The two LRMs we tested, o1-preview and o1-mini, are surprisingly effective, but this performance is still not robust, and degrades quickly with length.

analysis, we believe that o1’s architecture supplements an underlying LLM with System 2-like abilities, allowing it to outperform previous models.

As far as we can tell, o1 combines an underlying LLM, most likely a modified GPT-4o, into an RL-trained system that steers the creation, curation, and final selection of private Chain-of-Thought reasoning traces. Exact details are currently sparse, and so we can only speculate about its exact mechanisms. Our best guess is that there are two major differences between o1 and LLMs: an additional reinforcement learning pre-training phase (perhaps to learn the q-values of different CoTs from massive amounts of synthetic data) and a new adaptively scaling inference procedure (maybe it further refines learned q-values by something like rollout before selecting a particular CoT; see Appendix B. Regardless, what looks clear from the detail available is that this model is fundamentally different in nature from previous LLMs.

### 3.1 PLANNING

**Evaluating LRMs on PlanBench:** We test o1-preview and o1-mini on the original 600-instance PlanBench test set.<sup>5</sup> The full results can be seen in Table 2. These 600 Blocksworld instances range from three to five blocks, and require plans of between 2 to 16 steps to solve. Far surpassing any LLM, o1 correctly answers 97.8% of these instances. On Mystery Blocksworld, the model does not maintain this level of performance, but it does far surpass all previous models (which barely managed a few percent), answering 52.8% correctly. To test whether the exact obfuscation might be compromised because of data contamination, we also generated a new obfuscation using completely random strings, and presented these problems in a new, semantically equivalent prompt format with fully specified and unambiguous PDDL descriptions of both the domain and problem. This is presented in the table as Randomized Mystery Blocksworld. Exact prompts can be seen in the appendix. While performance did

<sup>5</sup>While for previous models, the model itself enforced the desired plan format, some modifications had to be made to accurately test o1’s abilities. In its current form, o1-preview does not always conform to explicit formatting restrictions. This is right in line with OpenAI’s injunction to keep o1 prompts "simple and direct" (OpenAI, 2024c). In order to extract the generated plans, we used GPT-4o-mini to translate them into PDDL, and wrote a small Python parser to strip any remaining extraneous symbols before evaluating each proposed plan.

Table 3: Rate of claiming that a problem is impossible by OpenAI’s o1-preview on 100 unsolvable and 600 solvable instances in the Blocksworld and Randomized Mystery Blocksworld domains. The True Negative rate is the percent of unsolvable instances that were *correctly* marked as unsolvable. The False Negative rate is the percent of solvable instances that were *incorrectly* marked as unsolvable. Previous models are not shown in this table as their true negative and false negative rates were generally 0% across the board.

Domain	Shots	o1-preview	
		True Negatives	False Negatives
Blocksworld	0-Shot	27%	0%
Randomized Mystery Blocksworld	1-Shot	16%	11.5%

dip further, 37.3% of instances are answered correctly, sharply contrasting the flat zeroes of previous models. The same pattern can be seen when evaluating Logistics and a freshly generated obfuscation of that domain. Despite the higher branching factor of the domain, o1-preview solves 94% of all 200 problems tested and achieves 52% on the obfuscated variant.

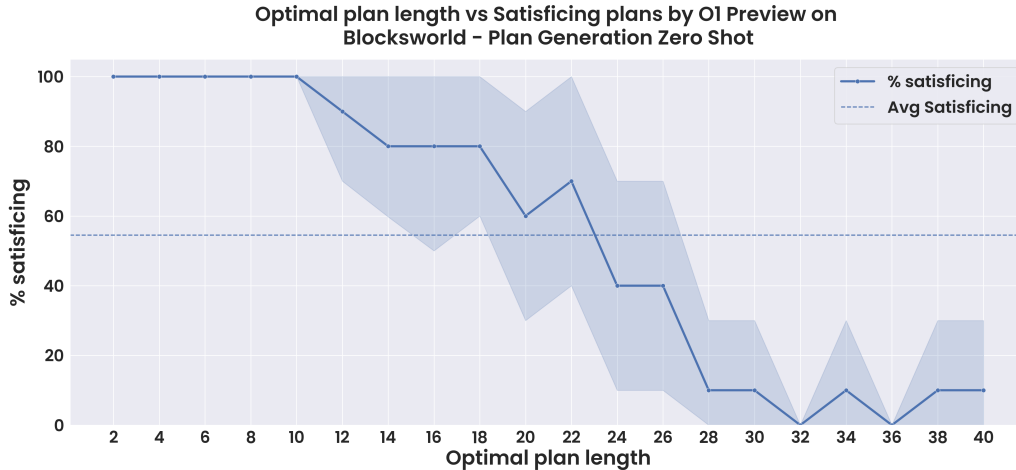


Figure 2: Extending even the (regular, not obfuscated) Blocksworld dataset to problems requiring greater numbers of steps worsens the performance of o1-preview. When tested on 110 instances which each require at least 20 steps to solve, it only manages 23.63%.

**Increasing Problem Size:** Standard LLM chain-of-thought prompting approaches are brittle, do not robustly scale with problem size, and fail to induce general algorithmic procedure-following (Stechly et al., 2024a). We extend planbench to a set of 110 harder Blocksworld problems. Problems in this set range from 6 to 20 blocks in length and require 20 to 40 step optimal plans. Without any obfuscation, we see performance quickly degrade from the 97.8% reported earlier. In fact, over these 110 instances, o1-preview only manages 23.63%, and most of this accuracy comes from correctly solving problems which require fewer than 28 steps. While these models are overall impressive, this shows that their performance is still far from robust. These results are collated together with a representative sampling of smaller instances in Figure 3, showing how performance begins to fall on problems requiring plans of more than 10 steps.

**Performance on Unsolvable Instances:** While planning problems normally require the agent to formulate a course of action to achieve a goal, an equally valid use of planning abilities is to recognize that a given goal *cannot* be accomplished by any plan. A real-world example of this is network vulnerability analysis, where an agent may wish to certify that

---

no plan of attack exists for a specified system (Boddy et al., 2005). So far, LLMs have struggled to recognize that some problems cannot be solved, instead confidently confabulating nonsensical answers. o1 was launched with the claim that it has started to overcome this issue, and can now accurately identify unsolvable problems (Brown, 2024). To test this systematically, we modified 100 instances from the easier three to five block test set by adding one `on(x,y)`-type conjunct to each instance’s goal state, making the goal unsatisfiable. We ensured our instances were unambiguous by giving the full PDDL representation of both the domain and the instance, to avoid quibbles such as "A is on B because A is on C and C is on B" where the model redefines the meanings of a potentially ambiguous natural language statement. The results are in Table 3. On Blocksworld, only 27% of all instances were correctly and explicitly identified by o1 as unsolvable. In 19% of all cases, the model returned a dot or some kind of "[empty plan]" marker, without any explanation or indication of unsolvability. We consider these incorrect, as "empty plan" is only the correct answer if the goal is already satisfied. In the remaining 54% of cases, the model generated a full (and therefore impossible and incorrect!) plan.

On Randomized Mystery Blocksworld, these numbers are worse: 16% of cases were correctly identified as unsolvable, 5% returned an empty plan, and the remaining 79% were answered with a full (impossible or goal-unsatisfying) plan. Therefore, unsolvable instances continue to be a problem for LRMs. Furthermore, this ability to sometimes note impossible plans correctly comes at a cost: now the model sometimes falsely claims that solvable problems are actually unsolvable. On Randomized Mystery Blocksworld, 11.5% of instances are incorrectly claimed to be impossible. These results can be seen in Table 3.

**o1’s Creative Justifications** While our main focus has been on providing a quantitative evaluation of o1’s performance on PlanBench, we have also noticed an o1 idiosyncrasy that is worth commenting on. When the model gives an incorrect answer, it also sometimes provides a creative, but nonsensical, justification for its decision. It is almost as if o1 has gone from *hallucinating* to *gaslighting*! In one case, it decided that an unsolvable problem is solvable because a goal condition, while not present in the final state, had been true for at some point during the execution, and thus should continue to count. In another, it declared that `on(a,c)` was true because, as it explained in a brief parenthetical, a was on b which was on c, and was thus a was somewhere above c, which should count as being "on top" of it. As we mentioned earlier, we changed our unsolvable instance prompts from natural language to PDDL in order to make it extremely clear that divergences from our exact definitions were disallowed.

**Extending to Harder Domains:** While STRIPS planning is in principle PSPACE-complete, the problems presented above are actually of a lower computational complexity. Plan existence for both Blocksworld and Logistics is polynomial (Hoffmann et al., 2006). Sokoban is a non-ergodic domain in which an agent moves around a constrained grid, pushing boxes by running into them, where the goal is to move every box to one of a set of final locations. It can be represented in PDDL form, and is thus amenable to classical planning techniques, and is known to be PSPACE-complete (Culberson, 1997).

Using a generator from the 2008 International Planning Competition (IPC, 1998), we generate 55 Sokoban instances with grid sizes ranging from  $4 \times 4$  to  $10 \times 10$ , 1 to 4 boxes, and 1 to 4 walls. We then create corresponding prompts in PDDL. o1-preview answers 12.7% of these instances correctly, while o1-mini is not far behind with 10.9%. For comparison, when evaluated on these same instances, Llama3.1-405B, despite doing the best of all LLMs on the Blocksworld sets, does not answer a single question correctly.

### 3.2 SCHEDULING

We also evaluate o1 on a set of scheduling problems that have previously been used in testing LLM capabilities.

**Graph Coloring:** We evaluated o1-mini on the set of 220 problems included in the codebase associated with Stechly et al. (2024b)’s paper, and found that it solved 96%, surpassing

Table 4: Cost per 100 instances (in USD). LRMs are significantly more expensive than LLMs.

Costs per 100 instances (in USD)									
Large Language Models							Large Reasoning Models		
Claude 3.5 (Sonnet)	Claude 3 (Opus)	GPT-4o	GPT-4o-mini	GPT-4	GPT-4 Turbo	Gemini 1.5 Pro	Gemini 1 Pro	o1-preview	o1-mini
\$0.44	\$1.70	\$0.65	\$0.02	\$1.80	\$1.20	\$0.33	\$0.03	\$42.12	\$3.69

the 16% reported by that paper for GPT-4. To test the full capabilities of the model, we extended the test set. Using the same Erdős–Rényi procedure with  $p = 0.4, n = 20$ , we generated 50 more graphs, and did not rejection sample for planarity. As in that paper, we precalculate the chromatic number and provide it in the prompt, asking that the model produce a coloring that uses exactly that number of colors. These harder graphs have 20 vertices and around 60 edges each. On this set, o1-mini solves 50%, while o1-preview solves 64%.

**Travel Planning:** We test o1 models on the 180 instance validation set of the sole-planning mode. In this mode, the model is provided upfront with all relevant information required to construct the requested itinerary. The previous state-of-the-art in direct prompting was 4.4%, achieved by GPT-4-turbo. o1-preview surpasses this, but only barely, solving only 10% of all instances. o1-mini does not beat even GPT-4-turbo, remaining at 1.67%.

**Natural Plan:** This benchmark consists of three domains: calendar scheduling, trip planning, and meeting planning. On calendar scheduling, o1-mini solves 94% of all instances correctly. Given cost constraints, we did not test o1-preview on this domain, as o1-mini tends to be a lower bound on its performance. Neither model performs well on trip planning—o1-preview reaches 4% and o1-mini only 1%, both lower than the previous state-of-the-art set by Gemini 1.5 Pro. When provided with prompts from the meeting planning domain, both models refuse to respond and flag the input as a potential terms violation (see Appendix E).

### 3.3 ACCURACY/COST TRADEOFFS AND GUARANTEES

With LRMs showing better performance on planning and scheduling problems, our evaluations must explicitly take into account the trade-offs that come from choosing general models over established deep and narrow systems. While o1-preview may provide higher accuracy than LLMs, it still fails to provide any correctness guarantees, and it is unclear that it is at all cost-effective. Unlike previous models, whose APIs only charge based on the number of input tokens and the number of output tokens (usually at a rate that is five times higher for the latter), o1’s price-per-call includes a surcharge based on the number of "reasoning tokens" it used—tokens generated as part of inference and not revealed to the user—which are charged at the significantly higher output token rate. Currently, end users have no control over the number of these tokens generated, a number which is expanded or limited by the model in its own opaque way. We have already run up a bill of over \$4000 for just the o1 model experiments reported in this evaluation!<sup>6</sup>)

Without exposing the ability to scale inference time to particular specifications, influence the internal ‘thinking’ process in task-specific ways, or ensure that intermediate steps are evaluated by trusted or sound verifiers, the o1 models are a coarse-grained choice in the space of cost, inference time, guarantees, and performance trade-offs. They aren’t, however, the only choices in that space, and reasonable LRM evaluations must take this into account (see similar arguments in (Katz et al., 2024; Kapoor et al., 2024)).

Classical planners like Fast Downward (Helmert, 2006) achieve 100% on our dataset in a fraction of the time, compute, and cost, while providing *guarantees that their answers are correct*. Running Fast Downward on a personal computer was essentially free in dollar terms

<sup>6</sup>The rich irony of researchers using tax payer provided research funds to pay private companies like OpenAI to evaluate their private commercial models is certainly not lost on us.



Table 5: Performance of LRM-modulo with o1-preview and o1-mini as the underlying LRMs on our hardest test sets: Blocksworld (20+ length plans), Sokoban, Graph Coloring (20 vertex), Travel Planning, and Trip Planning (10 cities). Due to cost constraints, we run each problem set for up to ten iterations, stopping early if the improvement per iteration levels off.

Total Instances	Domain	o1-preview			o1-mini		
		Direct	LRM Modulo	Iterations	Direct	LRM Modulo	Iterations
110	Blocksworld (hard)	23.65%	98.2%	7	0.90%	10%	4
55	Sokoban	12.70%	43.6%	7	10.90%	12.70%	4
50	Graph Coloring (hard)	64%	94%	10	50%	84%	15
180	Travel Planning	10%	65%	10	1.67%	41.11%	10
200	Trip Planning (10 cities)	4%	15.50%	4	1%	3.50%	4

and averaged 0.12 seconds per instance, which is many orders of magnitude faster than the average o1 clock times listed in table 2. It is also generally predictable, and can be scaled to harder instances very directly. Vanilla LLMs are typically very good at translating problems between formats, and could be used to do so in concert with a classical planner at a fraction of the cost of LRMs (e.g. Olmo et al. (2021); Liu et al. (2023)). For problems which don't have simple PDDL domain and instance specifications, LLM-Modulo systems may be a safer and cheaper approach: run a smaller, faster LLM in a loop with a sound verifier, so that the combined system will only output guaranteed correct solutions (e.g. Kambhampati et al. (2024); Romera-Paredes et al. (2024); Trinh et al. (2024)).

The correctness guarantees provided by these latter two methods are sorely lacking in LRMs like o1. A general reasoning system cannot be deployed in safety critical and non-ergodic domains if it continues to confidently make incorrect plans. o1 is a fully black box system, even more so than previous models, and OpenAI's decision to not only keep the architecture under wraps and hide the reasoning traces, but to warn away and even ban anyone who attempts to understand what is going on inside them (Edwards, 2024), makes interpretability nearly impossible, and reduces trust in the system overall.<sup>7</sup>

#### 4 LRM-MODULO TO IMPROVE O1 WITH GUARANTEES

We propose augmenting o1 with external verifiers to endow the combined system with soundness guarantees. While o1 is a stride in the direction of general-purpose, expressive planning systems, our results show that it cannot plan robustly when faced with harder instances, nor can it consistently recognize when instances are unsolvable, still providing incoherent plans in a majority of such cases. In other words, o1 is still fallible and without guarantees. Prior to the release of these models, the best way to coax planning capabilities out of LLMs has been to pair them with sound external verifier in generate-test frameworks, in what are known as LLM-Modulo systems (Kambhampati et al., 2024; Trinh et al., 2024). This framework is broadly applicable even beyond LLMs, and—given a sound verifier for some domain—requires only a generator expressive enough to provide guesses for that domain. Moreover, because of the built-in verification, it guarantees that any answer output is correct. For safety-critical systems, this is essential! High accuracies are not sufficient, especially when the underlying system—as is the case for both LLMs and even more so for LRMs—is an opaque black box. Therefore, we investigated integrating LRMs into LRM-modulo systems to both boost their overall performance and to provide much-needed guarantees over their outputs.

<sup>7</sup>The current model is also set to a default temperature of 1.0, which further reduces replicability and interpretability—for any given problem, it is never clear whether the result is merely the result of stochasticity. This compounds a problem with OpenAI models that has existed since at least GPT3. Temperature 0 never gave deterministic outputs, and worse, the logprobs provided by the OpenAI API for any given prompt have long been known to fluctuate wildly (Xuan, 2023).

---

486 Generate-test systems are limited by how good the generator is. A poor generator, such  
487 as one that produces completely random strings, may be capable of eventually producing  
488 the correct answer, but be so unlikely to do so at each iteration as to be useless, while an  
489 incomplete generator may never output the correct answer at all. LLMs and LRMs can be  
490 backprompted—that is, we can take feedback from the sound verifier and send it back to  
491 the model or modify the next prompt in some other way to increase the diversity of the  
492 responses generated—which may steer their next output towards the correct answer. Based  
493 on our results, o1 models are much better generators than anything that came before them,  
494 but, a priori, it is unclear if they are any more complete or capable of effectively utilizing  
495 backprompts or advice.

496 We test LRM-modulo setups on our five hardest test sets: 20+ length plan Blocksworld,  
497 Sokoban, 20 vertex graph coloring, OSU’s Travel Planning, and 10 city trip planning. Due  
498 to cost constraints, we limit the number of iterations to a maximum of ten, but we stop the  
499 system early once the performance increase from round to round has become mostly flat.<sup>8</sup>  
500 Even with so few iterations, we see significant jumps in performance across almost all of  
501 our domains. o1-preview’s performance on harder Blocksworld saturates within 7 iterations,  
502 with the combined system achieving 98.2%. Harder graph coloring shows similar results,  
503 going up to 94%. Perhaps most surprising, our most difficult domain, Sokoban, shows a  
504 significant jump from 12.7% to 43.6%. o1-mini-Modulo performance, while also impressive,  
505 only comes close on graph coloring and travel planning.

506 Our results seem to imply that these models are not only better generators, but also benefit  
507 more from the sound verification signal and provided feedback. However, we can’t really  
508 know *how* they use the critique provided—this question is likely crucial, but also unanswerable  
509 with OpenAI’s current stance against revealing the internal workings of the model or the  
510 intermediate reasoning tokens it generates.

511 With the high cost of o1 model queries, it is also crucial to examine in which situations these  
512 models are not just the best option, but the cost-effective one. Returning to LLM-modulo,  
513 where we use a smaller, cheaper, and faster LLM in a loop with a verifier, can provide similar  
514 or better performance in some domains. For example, in calendar scheduling, o1-mini costs  
515 \$2.70 to run over the entire test set, and has a final accuracy of 94%. GPT-4o-mini can be  
516 run for 50 iterations in a modulo loop for only \$2.48, beating that performance to get 97%,  
517 a figure that comes with guarantees not provided by just the base model.

## 518 5 CONCLUSION

519 In this paper, we investigated the performance of o1-preview and o1-mini—the new so-called  
520 LRMs—on a variety of planning and scheduling benchmarks. While LLMs have thus far  
521 failed to make much progress on the obfuscated (or "Mystery") versions of PlanBench  
522 domains, o1 shows the first bit of real progress. In general, it seems to have made impressive  
523 headway on benchmarks that were previously unassailable. However, when we evaluated  
524 the model on longer problems and on the question of determining solvability of potentially  
525 impossible instances, we found that these accuracy gains are not general nor robust. While  
526 o1 made some gains on scheduling problems, performing much better on graph coloring than  
527 previous models, these were not evenly distributed, only making some progress on OSU’s  
528 Travel Plan domain and the Natural Plan benchmark suite. We also discussed the critical  
529 accuracy/efficiency tradeoffs that are brought up by the fact that o1 that uses (and charges  
530 for) significant inference-time compute, as well as how it compares to other LLM-based  
531 approaches (such as LLM-Modulo (Kambhampati et al., 2024)) and dedicated solvers. Future  
532 evaluations will have to maintain a focus on these factors if they are to remain meaningful  
533 or relevant. Finally, we showed that approaches like LLM-Modulo (Kambhampati et al.,  
534 2024) can indeed be adapted to LRMs to further improve their performance and to provide  
535 much-needed guarantees. In essence, LRMs can replace LLMs as significantly better—but  
536 still fallible—generators in LLM-Modulo frameworks.

---

537  
538  
539 <sup>8</sup>In the case of trip planning, we stopped early because of mounting costs: for just 4 iterations,  
we incurred an API access cost of \$535!

- 
- 540 REFERENCES  
541  
542 Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David,  
543 Chelsea Finn, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, et al. Do as i can,  
544 not as i say: Grounding language in robotic affordances. *arXiv preprint arXiv:2204.01691*,  
545 2022.
- 546 Mark S Boddy, Johnathan Gohde, Thomas Haigh, and Steven A Harp. Course of action  
547 generation for cyber security using classical planning. In *ICAPS*, pp. 12–21, 2005.  
548
- 549 Noam Brown. o1 gets it right almost always, 2024. URL [https://x.com/polynoomial/  
550 status/1834280720493412724](https://x.com/polynoomial/status/1834280720493412724). Accessed: 2024-09-14.
- 551 Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz,  
552 Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. Sparks of artificial  
553 general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*, 2023.  
554
- 555 Jacques Carlier. The one-machine sequencing problem. *European Journal of Operational  
556 Research*, 11(1):42–47, 1982.
- 557 Joseph Culberson. Sokoban is pspace-complete. 1997.  
558
- 559 Nouha Dziri, Ximing Lu, Melanie Sclar, Xiang Lorraine Li, Liwei Jiang, Bill Yuchen Lin,  
560 Sean Welleck, Peter West, Chandra Bhagavatula, Ronan Le Bras, et al. Faith and fate:  
561 Limits of transformers on compositionality. *Advances in Neural Information Processing  
562 Systems*, 36, 2024.
- 563 Benj Edwards. Ban warnings fly as users dare to probe the  
564 “thoughts” of openai’s latest model. *Ars Technica*, September 2024.  
565 URL [https://arstechnica.com/information-technology/2024/09/  
566 openai-threatens-bans-for-probing-new-ai-models-reasoning-process/](https://arstechnica.com/information-technology/2024/09/openai-threatens-bans-for-probing-new-ai-models-reasoning-process/).
- 567 Malik Ghallab, Dana S. Nau, and Paolo Traverso. *Automated Plan-  
568 ning and Acting*. Cambridge University Press, 2016. ISBN 978-1-  
569 107-03727-4. URL [http://www.cambridge.org/de/academic/subjects/  
570 computer-science/artificial-intelligence-and-natural-language-processing/  
571 automated-planning-and-acting?format=HB](http://www.cambridge.org/de/academic/subjects/computer-science/artificial-intelligence-and-natural-language-processing/automated-planning-and-acting?format=HB).
- 572
- 573 Malte Helmert. The fast downward planning system. *Journal of Artificial Intelligence  
574 Research*, 26:191–246, 2006.
- 575
- 576 Jörg Hoffmann, Stefan Edelkamp, Sylvie Thiébaux, Roman Englert, Frederico Liporace,  
577 and Sebastian Trüg. Engineering benchmarks for planning: the domains used in the  
578 deterministic part of ipc-4. *Journal of Artificial Intelligence Research*, 26:453–541, 2006.
- 579 Wenlong Huang, Pieter Abbeel, Deepak Pathak, and Igor Mordatch. Language models as  
580 zero-shot planners: Extracting actionable knowledge for embodied agents. *arXiv preprint  
581 arXiv:2201.07207*, 2022a.
- 582
- 583 Wenlong Huang, Fei Xia, Ted Xiao, Harris Chan, Jacky Liang, Pete Florence, Andy Zeng,  
584 Jonathan Tompson, Igor Mordatch, Yevgen Chebotar, et al. Inner monologue: Embodied  
585 reasoning through planning with language models. *arXiv preprint arXiv:2207.05608*,  
586 2022b.
- 587 IPC. International planning competition. <https://www.icaps-conference.org/competitions>,  
588 1998.
- 589
- 590 Subbarao Kambhampati. Can large language models reason and plan? *Annals of the New  
591 York Academy of Sciences*, 1534(1):15–18, 2024.
- 592
- 593 Subbarao Kambhampati, Karthik Valmeekam, Lin Guan, Kaya Stechly, Mudit Verma,  
Siddhant Bhambri, Lucas Saldyt, and Anil Murthy. Llms can’t plan, but can help planning  
in llm-modulo frameworks. *arXiv preprint arXiv:2402.01817*, 2024.

---

594 Sayash Kapoor, Benedikt Stroebl, Zachary S. Siegel, Nitya Nadgir, and Arvind Narayanan.  
595 Ai agents that matter, 2024. URL <https://arxiv.org/abs/2407.01502>.  
596

597 Michael Katz, Harsha Kokel, Kavitha Srinivas, and Shirin Sohrabi. Thought of search:  
598 Planning with language models through the lens of efficiency, 2024. URL <https://arxiv.org/abs/2404.11833>.  
599

600 Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa.  
601 Large Language Models are Zero-Shot Reasoners. *arXiv preprint arXiv:2205.11916*, 2022.  
602

603 Bo Liu, Yuqian Jiang, Xiaohan Zhang, Qiang Liu, Shiqi Zhang, Joydeep Biswas, and Peter  
604 Stone. Llm+ p: Empowering large language models with optimal planning proficiency.  
605 *arXiv preprint arXiv:2304.11477*, 2023.  
606

607 Man Luo, Shrinidhi Kumbhar, Mihir Parmar, Neeraj Varshney, Pratyay Banerjee, Somak  
608 Aditya, Chitta Baral, et al. Towards logigluue: A brief survey and a benchmark for analyzing  
609 logical reasoning capabilities of language models. *arXiv preprint arXiv:2310.00836*, 2023.

610 Drew McDermott, Malik Ghallab, Adele E. Howe, Craig A. Knoblock, Ashwin Ram,  
611 Manuela M. Veloso, Daniel S. Weld, and David E. Wilkins. Pddl-the planning domain  
612 definition language. 1998.

613 Alberto Olmo, Sarath Sreedharan, and Subbarao Kambhampati. GPT3-to-plan: Extracting  
614 plans from text using GPT-3. *arXiv preprint arXiv:2106.07131*, 2021.  
615

616 OpenAI. Openai o1 system card. *preprint*, 2024a.  
617

618 OpenAI. Introducing openai o1-preview, 2024b. URL <https://openai.com/index/introducing-openai-o1-preview/>. Accessed: 2024-09-19.  
619

620 OpenAI. Guides: Reasoning, 2024c. URL <https://platform.openai.com/docs/guides/reasoning>. Accessed: 2024-09-14.  
621  
622

623 Jing Qian, Hong Wang, Zekun Li, Shiyang Li, and Xifeng Yan. Limitations of language  
624 models in arithmetic and symbolic induction. *arXiv preprint arXiv:2208.05051*, 2022.  
625

626 Manley Roberts, Himanshu Thakur, Christine Herlihy, Colin White, and Samuel Dooley.  
627 Data contamination through the lens of time. *arXiv preprint arXiv:2310.10628*, 2023.

628 Bernardino Romera-Paredes, Mohammadamin Barekatin, Alexander Novikov, Matej Balog,  
629 M Pawan Kumar, Emilien Dupont, Francisco JR Ruiz, Jordan S Ellenberg, Pengming  
630 Wang, Omar Fawzi, et al. Mathematical discoveries from program search with large  
631 language models. *Nature*, 625(7995):468–475, 2024.  
632

633 Sanjana Srivastava, Chengshu Li, Michael Lingelbach, Roberto Martín-Martín, Fei Xia,  
634 Kent Elliott Vainio, Zheng Lian, Cem Gokmen, Shyamal Buch, Karen Liu, et al. Behavior:  
635 Benchmark for everyday household activities in virtual, interactive, and ecological  
636 environments. In *Conference on robot learning*, pp. 477–490. PMLR, 2022.

637 Kaya Stechly, Karthik Valmeekam, and Subbarao Kambhampati. Chain of thoughtlessness:  
638 An analysis of cot in planning. *arXiv preprint arXiv:2405.04776*, 2024a.  
639

640 Kaya Stechly, Karthik Valmeekam, and Subbarao Kambhampati. On the self-verification  
641 limitations of large language models on reasoning and planning tasks. *arXiv preprint*  
642 *arXiv:2402.08115*, 2024b.

643 Trieu H Trinh, Yuhuai Wu, Quoc V Le, He He, and Thang Luong. Solving olympiad geometry  
644 without human demonstrations. *Nature*, 625(7995):476–482, 2024.  
645

646 Karthik Valmeekam, Matthew Marquez, Sarath Sreedharan, and Subbarao Kambhampati.  
647 On the planning abilities of large language models—a critical investigation. *arXiv preprint*  
*arXiv:2305.15771*, 2023.

---

648 Karthik Valmeekam, Matthew Marquez, Alberto Olmo, Sarath Sreedharan, and Subbarao  
649 Kambhampati. Planbench: An extensible benchmark for evaluating large language models  
650 on planning and reasoning about change. *Advances in Neural Information Processing*  
651 *Systems*, 36, 2024.

652 Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Chi, Quoc Le, and Denny  
653 Zhou. Chain of thought prompting elicits reasoning in large language models. *arXiv*  
654 *preprint arXiv:2201.11903*, 2022.

655 Jian Xie, Kai Zhang, Jiangjie Chen, Tinghui Zhu, Renze Lou, Yuandong Tian, Yanghua  
656 Xiao, and Yu Su. Travelplanner: A benchmark for real-world planning with language  
657 agents. *arXiv preprint arXiv:2402.01622*, 2024.

658 Tan Zhi Xuan. log. probs returned by openai’s api are \*incredibly\* unstable, 2023. URL  
659 <https://x.com/xuanalogue/status/1653280462935146496>. Accessed: 2024-09-14.

660 Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R Narasimhan, and  
661 Yuan Cao. React: Synergizing reasoning and acting in language models. In *The Eleventh*  
662 *International Conference on Learning Representations*, 2023. URL [https://openreview.](https://openreview.net/forum?id=WE_vluYUL-X)  
663 [net/forum?id=WE\\_vluYUL-X](https://openreview.net/forum?id=WE_vluYUL-X).

664 Huaixiu Steven Zheng, Swaroop Mishra, Hugh Zhang, Xinyun Chen, Minmin Chen, Azade  
665 Nova, Le Hou, Heng-Tze Cheng, Quoc V Le, Ed H Chi, et al. Natural plan: Benchmarking  
666 llms on natural language planning. *arXiv preprint arXiv:2406.04520*, 2024.

667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701

---

## APPENDIX

### A FURTHER DISCUSSION OF LLM PLANNING PERFORMANCE

LLMs are highly capable at providing translations between equivalent representations (Olmo et al., 2021). This fact, combined with their significantly higher performance on the unobfuscated version of the Blocksworld domain, predicts that—if they are capable of composing reasoning operations—the performance gap between Mystery Blocksworld and classic Blocksworld should shrink substantially if the translation from Mystery Blocksworld back into Blocksworld is explicitly provided. However, when we provide this in the prompt (see Appendix G), performance only improves a very small amount: GPT-4 achieves 10%.

### B SPECULATIONS ABOUT O1 INTERNAL OPERATION

While our evaluation of o1 did not depend on any specific assumption about its operation, we did have a working model of o1 based on the very skimpy description that was provided in the blog post that accompanies o1’s release (OpenAI, 2024b). Verifying our model is unfortunately made infeasible by the fact that o1 doesn’t actually provide any trace of its operations (even during the costly inference stage), and OpenAI warns that API access will be revoked if any attempts are made to surface its reasoning tokens.

There are two things—“reinforcement learning” and “Private Chain-of-Thought (CoT)” that are mentioned in the writeup. So imagine you are trying to transplant a “generalized AlphaGo”—let’s call it GPTGo—onto the underlying LLM token prediction substrate.

To do this, you need to know

1. What are the GPTGo moves? For AlphaGo, we had GO moves). What would be the right moves when the task is just “complete the prompt the right way”?
2. Where is it getting its external success/failure signal from? For AlphaGo, we had simulators/verifiers giving the success/failure signal. The most interesting question in transplanting the self-play idea to a general AI agent is where is it getting this signal?

Our guess is that the moves are auto-generated CoTs (thus the moves have a very high branching factor). Let’s assume—for simplification—that we have a CoT-generating LLM, that generates these CoTs conditioned on the prompt. (It is not clear if the CoT’s are domain independent of the “think step by step” variety (Kojima et al., 2022) or domain/task specific, or a combination.)

The success signal is likely from massive amounts of synthetic training data with correct answers. When the completed prompt is seen to contain the correct answer (presumably judged by the LLM itself), then the episode is considered a success, and a failure otherwise.

The task for the reinforcement learner then is: Given the original problem prompt, generate and select a CoT, and use it to continue to extend the prompt (possibly generating subgoal CoTs after every few stages). Get the final success/failure signal for the example (for which you do have answer).

The RL stage may involve training on a a huge number of training examples with answers. The training examples with answers can either be coming from benchmarks, or from synthetic data with problems and their solutions—using external solvers. In this phase the RL part attempts to learn the q-values of the CoT moves (much like AlphaGo learns the q-values of the moves of the Go). (The q-values learning may be incorporated into the internal weights of the CoT generator LLM). At this point, we have a CoT move generator that is better than the random one before the RL stage

During the inference stage—which OpenAI says can be indefinitely long (although it is currently capped internally by them, with no external control), like AlphaGo, o1 might be further improving its evaluation of the q-values of the CoT moves in the context of the current prompt. While AlphaGo used MCT-based rollouts, we obviously don’t know the

756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809

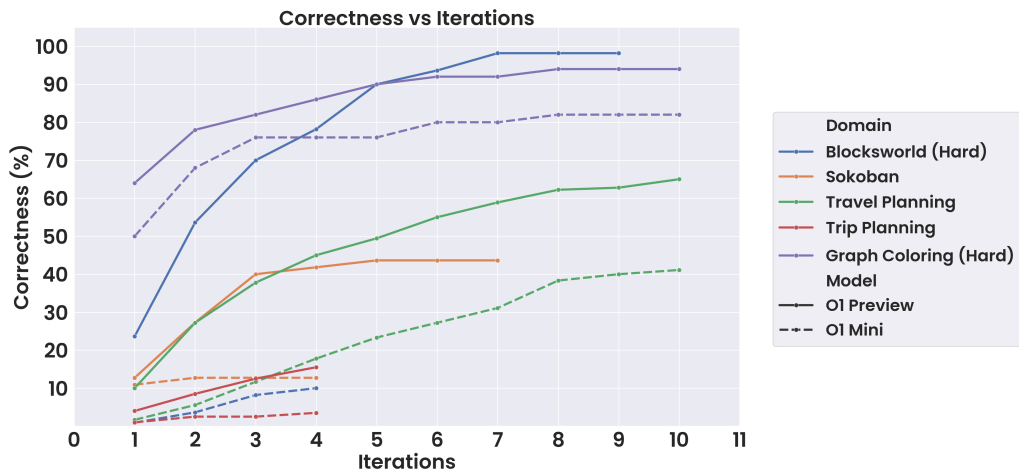


Figure 3: LRM-Modulo significantly improves performance over direct prompting as we increase the number of iterations.

mechanism o1 uses. The announcement only says that at inference stage a long chain of thought is added to the original prompt (and o1 does charge the end users for its “reasoning tokens,” which are never seen by the end user, at the same high rate as the output tokens). In this sense, our speculations seem to be consistent, even though it is not clear whether the reasoning tokens are proportional to the entire inference-stage computation, or just represent the final sequence of CoT moves that get selected after the rollout-like inference stage.

Some corollaries of our speculation are:

1. Note that this use of RL is very different from that in RLHF, which can be seen as a fine tuning stage for an LLM that keeps the inference step unchanged. It is also different from techniques, including OpenAI’s, that advocated fine tuning both on synthetic data accompanied with derivational traces—these too will be a form of finetuning that leave inference stage unchanged. Here o1 could, in theory, be getting significantly more leverage out of the data by learning move (auto CoT) generators.
2. There still are no guarantees that the answers provided are “correct”—they may be probabilistically a little more correct (subject to the training data). If you want guarantees, you still will need some sort of LLM-Modulo approach even on top of this.
3. It is certainly not clear that anyone will be willing to really wait for long periods of time during inference (it is already painful to wait for 10 sec for a 10 word last letter concatenation!). The kind of people who will wait for longer periods would certainly want guarantees—and there are deep and narrow System 2’s a plenty that can be used for many such cases.
4. There is a bit of a *Ship of Theseus* feel to calling o1 an LLM—considering how far it is from the other LLM models (all of which essentially have teacher-forced training and sub-real-time next token prediction. That said, this is certainly an interesting way to build a generalized system 2’ish component on top of LLM substrates—but without guarantees.

### C LRM-MODULO PERFORMANCE OVER ITERATIONS

See Figure 3

### D GEMINI 1.5 PRO RESPONSE TO MYSTERY BLOCKSWORLD

finish\_reason: SAFETY

```

810
811 safety_ratings {
812   category: HARM_CATEGORY_SEXUALLY_EXPLICIT
813   probability: NEGLIGIBLE
814 }
815 safety_ratings {
816   category: HARM_CATEGORY_HATE_SPEECH
817   probability: NEGLIGIBLE
818 }
819 safety_ratings {
820   category: HARM_CATEGORY_HARASSMENT
821   probability: NEGLIGIBLE
822 }
823 safety_ratings {
824   category: HARM_CATEGORY_DANGEROUS_CONTENT
825   probability: MEDIUM
826 }

```

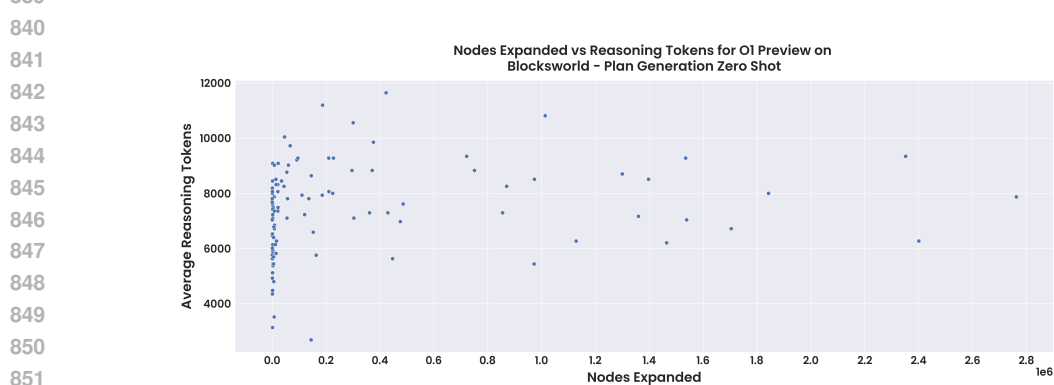
## E O1-PREVIEW AND O1-MINI RESPONSE TO MEETING PLANNING

```

827
828
829 openai.BadRequestError: b'{'
830   "error": {
831     "message": "Invalid prompt: your prompt was flagged as potentially violating
832               our usage policy. Please try again with a different prompt.",
833     "type": "invalid_request_error",
834     "param": null,
835     "code": "invalid_prompt"
836   }
837 }'

```

## F O1 TOKEN USE VERSUS PROBLEM DIFFICULTY



852  
853 Figure 4: The number of reasoning tokens used by o1-preview when solving Blocksworld  
854 instances does not track the number of nodes that need to be expanded to solve the problem.  
855

## 856 G PROMPT TO TRANSLATE FROM MYSTERY BACK TO BLOCKSWORLD

```

857
858 I am playing with a set of objects. Here are the actions I can do
859
860   Attack object
861   Feast object from another object
862   Succumb object
863   Overcome object from another object
864
865 I have the following restrictions on my actions:

```



---

864 To perform Attack action, the following facts need to be true: Province object,  
865 Planet object, Harmony.  
866 Once Attack action is performed the following facts will be true: Pain object.  
867 Once Attack action is performed the following facts will be false: Province object,  
868 Planet object, Harmony.  
869 To perform Succumb action, the following facts need to be true: Pain object.  
870 Once Succumb action is performed the following facts will be true: Province object,  
871 Planet object, Harmony.  
872 Once Succumb action is performed the following facts will be false: Pain object.  
873 To perform Overcome action, the following needs to be true: Province other object,  
874 Pain object.  
875 Once Overcome action is performed the following will be true: Harmony, Province  
876 object, Object Craves other object.  
877 Once Overcome action is performed the following will be false: Province other object  
878 , Pain object.  
879 To perform Feast action, the following needs to be true: Object Craves other object,  
880 Province object, Harmony.  
881 Once Feast action is performed the following will be true: Pain object, Province  
882 other object.  
883 Once Feast action is performed the following will be false:, Object Craves other  
884 object, Province object, Harmony.

883 You will be given a set of initial conditions and a goal condition. To solve the  
884 problem, you will have to tell me which actions to take and in which order in  
885 order to achieve the goal.  
886

887 Please provide your answers using the above terminology. However, you may find it  
888 helpful to translate the above description into a common-sense format while  
889 working out your solution. Just remember to translate it back later!  
890 Instead of thinking in terms of "objects", think in terms of different alphabet  
891 blocks (block A, block B, etc.) which you are stacking (using just one hand) in  
892 towers on a table.

893 Then the "facts" that are true or false at a given time are really facts about the  
894 blocks and the hand:

895 "Province object a" just means that "block A is clear" or, equivalently, "  
896 nothing is on top of block A"  
897 "Planet object a" is another way of saying "block A is on the table"  
898 "Harmony" is a codeword for "my hand isn't holding anything"  
899 "Pain object a" = "the hand is holding block A"  
"object a Craves object b" translates to "block A is on top of block B"

900 And the "actions" can be seen as stacking and unstacking of blocks (where the  
901 restrictions stop us from picking up the bottom block in a tower or holding  
902 more than one block in the hand at a time):

903 "Attack object a" translates to "pick up block A directly from the table"  
904 "Feast object a from object b" translates to "pick up block A from directly on  
905 top of block B"  
906 "Succumb object a" translates to "put block A directly on the table"  
907 "Overcome object a from object b" translates to "put block A directly on top  
908 of block B"

909 If you do use this framing, make sure to remember to translate back into the  
910 original terms.

911 H PROMPTS FOR BLOCKSWORLD

912 H.1 (SOLVABLE) BLOCKSWORLD INSTANCES - ZERO-SHOT IN NATURAL LANGUAGE

915 I am playing with a set of blocks where I need to arrange the blocks into stacks.  
916 Here are the actions I can do

917 Pick up a block

---

918 Unstack a block from on top of another block  
919 Put down a block  
920 Stack a block on top of another block  
921  
922 I have the following restrictions on my actions:  
923 I can only pick up or unstack one block at a time.  
924 I can only pick up or unstack a block if my hand is empty.  
925 I can only pick up a block if the block is on the table and the block is clear. A  
926 block is clear if the block has no other blocks on top of it and if the block  
927 is not picked up.  
928 I can only unstack a block from on top of another block if the block I am unstacking  
929 was really on top of the other block.  
930 I can only unstack a block from on top of another block if the block I am unstacking  
931 is clear.  
932 Once I pick up or unstack a block, I am holding the block.  
933 I can only put down a block that I am holding.  
934 I can only stack a block on top of another block if I am holding the block being  
935 stacked.  
936 I can only stack a block on top of another block if the block onto which I am  
937 stacking the block is clear.  
938 Once I put down or stack a block, my hand becomes empty.  
939 Once you stack a block on top of a second block, the second block is no longer clear  
940 .  
941  
942 [STATEMENT]  
943 As initial conditions I have that, the red block is clear, the blue block is clear,  
944 the yellow block is clear, the hand is empty, the blue block is on top of the  
945 orange block, the red block is on the table, the orange block is on the table  
946 and the yellow block is on the table.  
947 My goal is to have that the orange block is on top of the blue block.  
948  
949 What is the plan to achieve my goal? Just give the actions in the plan.  
950  
951  
952 =====o1-preview RESPONSE=====

953 To achieve your goal of stacking the orange block on top of the blue block, you can  
954 follow these steps:

- 955 1. **\*\*Unstack the blue block from on top of the orange block.\*\***  
956 - Since the blue block is clear and your hand is empty, you can unstack it from  
957 the orange block.
- 958 2. **\*\*Put down the blue block on the table.\*\***  
959 - With the blue block in your hand, you can place it onto the table.
- 960 3. **\*\*Pick up the orange block from the table.\*\***  
961 - The orange block is now clear (no blocks on top of it), and your hand is empty  
962 after putting down the blue block, so you can pick it up.
- 963 4. **\*\*Stack the orange block on top of the blue block.\*\***  
964 - The blue block is clear on the table, so you can stack the orange block you're  
965 holding onto it.

966 **\*\*Actions in the plan:\*\***

- 967 - **\*\*Unstack\*\*** the blue block from the orange block.
- 968 - **\*\*Put down\*\*** the blue block on the table.
- 969 - **\*\*Pick up\*\*** the orange block from the table.
- 970 - **\*\*Stack\*\*** the orange block on top of the blue block.

971  
972  
973 H.2 (SOLVABLE) HARDER BLOCKSWORLD INSTANCES - ZERO-SHOT IN PDDL

974 The following is a PDDL specification of a planning problem. The first part, under  
975 the heading [DOMAIN], is the domain file. The second part, under the heading [  
976 QUERY PROBLEM], is the problem file. Using this information, which is correct,  
977 and no further assumptions, find a plan which, when run from the specified  
978 initial state, satisfies the specified goal. Provide your answer as a sequence

---

```

972         of actions in PDDL format. An action ACTION which acts on two objects OBJ1 and
973         OBJ2 would be written (ACTION OBJ1 OBJ2). Do not provide anything else in your
974         answer.
975 [DOMAIN]
976 (define (domain blocksworld-4ops)
977   (:requirements :strips)
978   (:predicates (clear ?x)
979                 (ontable ?x)
980                 (handempty)
981                 (holding ?x)
982                 (on ?x ?y))
983
984   (:action pick-up
985     :parameters (?ob)
986     :precondition (and (clear ?ob) (ontable ?ob) (handempty))
987     :effect (and (holding ?ob) (not (clear ?ob)) (not (ontable ?ob))
988                (not (handempty))))
989
990   (:action put-down
991     :parameters (?ob)
992     :precondition (holding ?ob)
993     :effect (and (clear ?ob) (handempty) (ontable ?ob)
994                (not (holding ?ob))))
995
996   (:action stack
997     :parameters (?ob ?underob)
998     :precondition (and (clear ?underob) (holding ?ob))
999     :effect (and (handempty) (clear ?ob) (on ?ob ?underob)
1000                (not (clear ?underob)) (not (holding ?ob))))
1001
1002   (:action unstack
1003     :parameters (?ob ?underob)
1004     :precondition (and (on ?ob ?underob) (clear ?ob) (handempty))
1005     :effect (and (holding ?ob) (clear ?underob)
1006                (not (on ?ob ?underob)) (not (clear ?ob)) (not (handempty))))))
1007
1008 [QUERY PROBLEM]
1009 (define(problem BW-rand-6)
1010   (:domain blocksworld-4ops)
1011   (:objects a b c d e f )
1012   (:init
1013     (handempty)
1014     (ontable a)
1015     (on b f)
1016     (on c a)
1017     (on d b)
1018     (on e c)
1019     (on f e)
1020     (clear d)
1021   )
1022   (:goal
1023     (and
1024       (on a f)
1025       (on b c)
1026       (on c a)
1027       (on d b)
1028       (on f e))
1029   )
1030   )
1031
1032 [PLAN]
1033
1034 =====o1-preview RESPONSE=====
1035 (unstack d b)

```

---

```

1026 (put-down d)
1027 (unstack b f)
1028 (stack b c)
1029 (pick-up d)
1030 (stack d b)
1031 (unstack f e)
1032 (put-down f)
1033 (unstack e c)
1034 (put-down e)
1035 (pick-up f)
1036 (stack f e)
1037 (unstack c a)
1038 (put-down c)
1039 (pick-up a)
1040 (stack a f)
1041 (pick-up c)
1042 (stack c a)
1043 (pick-up b)
1044 (stack b c)
1045 (pick-up d)
1046 (stack d b)

```

### 1045 H.3 (SOLVABLE) HARDER BLOCKSWORLD INSTANCES - BACKPROMPT

```

1047 The above plan is invalid. The following action at step 4 has an unmet precondition:
1048 stack b c
1049 The unmet precondition is:
1050 (clear c)
1051 Provide your revised plan as a sequence of actions in PDDL format. An action ACTION
1052 which acts on two objects OBJ1 and OBJ2 would be written (ACTION OBJ1 OBJ2). Do
1053 not provide anything else in your answer.

```

```
1054 [PLAN]
```

### 1056 H.4 UNSOLVABLE BLOCKSWORLD INSTANCES - ZERO-SHOT IN PDDL

```

1058 The following is a PDDL specification of a planning problem. The first part, under
1059 the heading [DOMAIN], is the domain file. The second part, under the heading [
1060 QUERY PROBLEM], is the problem file. Using this information, which is correct,
1061 and no further assumptions, find a plan which, when run from the specified
1062 initial state, satisfies the specified goal. Provide your answer as a sequence
1063 of actions in PDDL format. An action ACTION which acts on two objects OBJ1 and
1064 OBJ2 would be written (ACTION OBJ1 OBJ2). Do not provide anything else in your
1065 answer.

```

```

1065 [DOMAIN]
1066 (define (domain blocksworld-4ops)
1067   (:requirements :strips)
1068   (:predicates (clear ?x)
1069                (ontable ?x)
1070                (handempty)
1071                (holding ?x)
1072                (on ?x ?y))
1073   (:action pick-up
1074     :parameters (?ob)
1075     :precondition (and (clear ?ob) (ontable ?ob) (handempty))
1076     :effect (and (holding ?ob) (not (clear ?ob)) (not (ontable ?ob))
1077                 (not (handempty))))
1078   (:action put-down
1079     :parameters (?ob)
1080     :precondition (holding ?ob)
1081     :effect (and (clear ?ob) (handempty) (ontable ?ob)

```

```

1080         (not (holding ?ob))))
1081
1082 (:action stack
1083  :parameters (?ob ?underob)
1084  :precondition (and (clear ?underob) (holding ?ob))
1085  :effect (and (handempty) (clear ?ob) (on ?ob ?underob)
1086            (not (clear ?underob)) (not (holding ?ob))))
1087
1087 (:action unstack
1088  :parameters (?ob ?underob)
1089  :precondition (and (on ?ob ?underob) (clear ?ob) (handempty))
1090  :effect (and (holding ?ob) (clear ?underob)
1091            (not (on ?ob ?underob)) (not (clear ?ob)) (not (handempty))))
1092
1092 [QUERY PROBLEM]
1093 (define(problem BW-rand-4)
1094  (:domain blocksworld-4ops)
1095  (:objects a b c d )
1096  (:init
1097   (handempty)
1098   (on a b)
1099   (on b c)
1100   (on c d)
1101   (ontable d)
1102   (clear a)
1103   )
1104  (:goal
1105   (and
1106     (on d c)
1107     (on c d)
1108   )
1109  )
1110  )
1111
1111 [PLAN]
1112 =====o1-preview RESPONSE=====
1113 **No plan exists that satisfies the goal with the given initial state and domain.**
1114
1115 I PROMPTS FOR MYSTERY BLOCKSWORLD
1116
1117 I.1 MYSTERY BLOCKSWORLD INSTANCES - ZERO-SHOT IN NATURAL LANGUAGE
1118
1119 I am playing with a set of objects. Here are the actions I can do
1120
1121   Attack object
1122   Feast object from another object
1123   Succumb object
1124   Overcome object from another object
1125
1126 I have the following restrictions on my actions:
1127   To perform Attack action, the following facts need to be true: Province object,
1128   Planet object, Harmony.
1129   Once Attack action is performed the following facts will be true: Pain object.
1130   Once Attack action is performed the following facts will be false: Province
1131   object, Planet object, Harmony.
1132   To perform Succumb action, the following facts need to be true: Pain object.
1133   Once Succumb action is performed the following facts will be true: Province
1134   object, Planet object, Harmony.
1135   Once Succumb action is performed the following facts will be false: Pain object.
1136   To perform Overcome action, the following needs to be true: Province other
1137   object, Pain object.
1138   Once Overcome action is performed the following will be true: Harmony, Province
1139   object, Object Craves other object.

```

---

1134           Once Overcome action is performed the following will be false: Province other  
1135           object, Pain object.  
1136           To perform Feast action, the following needs to be true: Object Craves other  
1137           object, Province object, Harmony.  
1138           Once Feast action is performed the following will be true: Pain object, Province  
1139           other object.  
1140           Once Feast action is performed the following will be false:, Object Craves other  
1141           object, Province object, Harmony.

1142 [STATEMENT]  
1143 As initial conditions I have that, object b craves object c, harmony, planet object  
1144           a, planet object c, planet object d, province object a, province object b and  
1145           province object d.  
1146 My goal is to have that object c craves object b.

1147 What is the plan to achieve my goal? Just give the actions in the plan.  
1148 =====o1-preview RESPONSE=====

1149 To achieve the goal where object c craves object b, you can perform the following  
1150           sequence of actions:

1151 1. **\*\*Feast object b from object c\*\***  
1152     - Preconditions:  
1153       - 'Object b craves object c' (initial fact)  
1154       - 'Province object b' (initial fact)  
1155       - 'Harmony' (initial fact)  
1156     - Effects:  
1157       - Adds: 'Pain object b', 'Province object c'  
1158       - Deletes: 'Object b craves object c', 'Province object b', 'Harmony'

1159 2. **\*\*Succumb object b\*\***  
1160     - Preconditions:  
1161       - 'Pain object b' (from previous action)  
1162     - Effects:  
1163       - Adds: 'Province object b', 'Planet object b', 'Harmony'  
1164       - Deletes: 'Pain object b'

1165 3. **\*\*Attack object c\*\***  
1166     - Preconditions:  
1167       - 'Province object c' (from step 1)  
1168       - 'Planet object c' (initial fact)  
1169       - 'Harmony' (restored in step 2)  
1170     - Effects:  
1171       - Adds: 'Pain object c'  
1172       - Deletes: 'Province object c', 'Planet object c', 'Harmony'

1173 4. **\*\*Overcome object c from object b\*\***  
1174     - Preconditions:  
1175       - 'Province object b' (restored in step 2)  
1176       - 'Pain object c' (from step 3)  
1177     - Effects:  
1178       - Adds: 'Harmony', 'Province object c', 'Object c craves object b' (goal  
1179           achieved)  
1180       - Deletes: 'Province object b', 'Pain object c'

1181 By following this plan, you manipulate the states of the objects through the actions  
1182           provided, ultimately achieving the desired state where 'Object c craves object  
1183           b'.  
1184 **\*\*Answer:\*\***  
1185 To achieve the goal, perform these actions in order:

1186 1. Feast object b from object c  
1187 2. Succumb object b  
1188 3. Attack object c  
1189 4. Overcome object c from object b

1188 J PROMPTS FOR RANDOMIZED MYSTERY BLOCKSWORLD  
1189  
1190 J.1 RANDOMIZED MYSTERY BLOCKSWORLD INSTANCES - ZERO-SHOT IN NATURAL  
1191 LANGUAGE  
1192  
1193 I am playing with a set of objects. Here are the actions I can do  
1194  
1195 J4gv801gnu2it0yj object\_0 object\_1.  
1196 U64y1a9apusmslxb object\_0.  
1197 Kip9uw781pv62umn object\_0.  
1198 Wio5amhq7814n006 object\_0 object\_1.  
1199  
1200 I have the following restrictions on my actions:  
1201 To perform j4gv801gnu2it0yj action, the following facts need to be true:  
1202 tv30k33pzoulql6w object\_1, and a46zhwykn6jvbabk object\_0  
1203 Once j4gv801gnu2it0yj is performed the following facts will be true:  
1204 cql4o62plyeke3ok, tv30k33pzoulql6w object\_0, and b6e9q4r60gagvdcn object\_0  
1205 object\_1  
1206 Once j4gv801gnu2it0yj is performed the following facts will be false:  
1207 tv30k33pzoulql6w object\_1, and a46zhwykn6jvbabk object\_0  
1208 To perform u64y1a9apusmslxb action, the following facts need to be true:  
1209 tv30k33pzoulql6w object\_0, af9tse23ljclsqad object\_0, and cql4o62plyeke3ok  
1210 Once u64y1a9apusmslxb is performed the following facts will be true:  
1211 a46zhwykn6jvbabk object\_0  
1212 Once u64y1a9apusmslxb is performed the following facts will be false:  
1213 tv30k33pzoulql6w object\_0, af9tse23ljclsqad object\_0, and cql4o62plyeke3ok  
1214 To perform kip9uw781pv62umn action, the following facts need to be true:  
1215 a46zhwykn6jvbabk object\_0  
1216 Once kip9uw781pv62umn is performed the following facts will be true:  
1217 tv30k33pzoulql6w object\_0, cql4o62plyeke3ok, and af9tse23ljclsqad object\_0  
1218 Once kip9uw781pv62umn is performed the following facts will be false:  
1219 a46zhwykn6jvbabk object\_0  
1220 To perform wio5amhq7814n006 action, the following facts need to be true:  
1221 b6e9q4r60gagvdcn object\_0 object\_1, tv30k33pzoulql6w object\_0, and  
1222 cql4o62plyeke3ok  
1223 Once wio5amhq7814n006 is performed the following facts will be true:  
1224 a46zhwykn6jvbabk object\_0, and tv30k33pzoulql6w object\_1  
1225 Once wio5amhq7814n006 is performed the following facts will be false:  
1226 b6e9q4r60gagvdcn object\_0 object\_1, tv30k33pzoulql6w object\_0, and  
1227 cql4o62plyeke3ok  
1228  
1229 [STATEMENT]  
1230 As initial conditions I have that, af9tse23ljclsqad object\_0, af9tse23ljclsqad  
1231 object\_2, af9tse23ljclsqad object\_3, b6e9q4r60gagvdcn object\_1 object\_2,  
1232 cql4o62plyeke3ok, tv30k33pzoulql6w object\_0, tv30k33pzoulql6w object\_1 and  
1233 tv30k33pzoulql6w object\_3.  
1234 My goal is to have that b6e9q4r60gagvdcn object\_2 object\_1.  
1235  
1236 What is the plan to achieve my goal? Just give the actions in the plan.  
1237 =====o1-preview RESPONSE=====

1238 To achieve the goal 'b6e9q4r60gagvdcn object\_2 object\_1', you need to perform the  
1239 following sequence of actions:

- 1240 1. \*\*wio5amhq7814n006 object\_1 object\_2\*\*  
1241 - This action reverses the relationship between 'object\_1' and 'object\_2',  
1242 setting up the necessary conditions for the next steps.
- 1243 2. \*\*kip9uw781pv62umn object\_1\*\*  
1244 - By performing this action on 'object\_1', you restore certain facts ('  
1245 cql4o62plyeke3ok') required for subsequent actions and adjust the state of '  
1246 object\_1'.
- 1247 3. \*\*u64y1a9apusmslxb object\_2\*\*  
1248 - This action modifies 'object\_2' to have the required property 'a46zhwykn6jvbabk  
1249 object\_2', which is a precondition for the final action.
- 1250 4. \*\*j4gv801gnu2it0yj object\_2 object\_1\*\*

---

1242 - Performing this action with 'object\_2' and 'object\_1' establishes the goal fact  
1243 'b6e9q4r60gagvdcn object\_2 object\_1'.  
1244

1245 **\*\*Plan Summary:\*\***

1246 - **\*\*Action 1:\*\*** 'wio5amhq7814n006 object\_1 object\_2'  
1247 - **\*\*Action 2:\*\*** 'kip9uw781pv62umn object\_1'  
1248 - **\*\*Action 3:\*\*** 'u64y1a9apusmslxb object\_2'  
1249 - **\*\*Action 4:\*\*** 'j4gv801gnu2it0yj object\_2 object\_1'

1250  
1251 This sequence of actions satisfies all the necessary preconditions step-by-step and  
1252 ultimately achieves your goal.  
1253

1254 J.2 UNSOLVABLE RANDOMIZED MYSTERY BLOCKSWORLD INSTANCES - ZERO-SHOT IN  
1255 PDDL  
1256

1257 The following is a PDDL specification of a planning problem. The first part, under  
1258 the heading [DOMAIN], is the domain file. The second part, under the heading [  
1259 QUERY PROBLEM], is the problem file. Using this information, which is correct,  
1260 and no further assumptions, find a plan which, when run from the specified  
1261 initial state, satisfies the specified goal. Provide your answer as a sequence  
1262 of actions in PDDL format. An action ACTION which acts on two objects OBJ1 and  
1263 OBJ2 would be written (ACTION OBJ1 OBJ2). Do not provide anything else in your  
1264 answer.

1265 [DOMAIN]  
1266 (define (domain xaji0y)  
1267 (:requirements :equality)  
1268 (:predicates  
1269 (tv30k33pzoulql6w ?x1 - object)  
1270 (af9tse23ljclsqad ?x1 - object)  
1271 (cql4o62p1yeke3ok )  
1272 (a46zhwykn6jvbabk ?x1 - object)  
1273 (b6e9q4r60gagvdcn ?x1 - object ?x2 - object)  
1274 )  
1275 (:action u64y1a9apusmslxb  
1276 :parameters (?ob1 - object)  
1277 :precondition (and (tv30k33pzoulql6w ?ob1) (af9tse23ljclsqad ?ob1) (  
1278 cql4o62p1yeke3ok ))  
1279 :effect (and  
1280 (a46zhwykn6jvbabk ?ob1)  
1281 (not (tv30k33pzoulql6w ?ob1))  
1282 (not (af9tse23ljclsqad ?ob1))  
1283 (not (cql4o62p1yeke3ok )))  
1284 )  
1285 (:action kip9uw781pv62umn  
1286 :parameters (?ob1 - object)  
1287 :precondition (a46zhwykn6jvbabk ?ob1)  
1288 :effect (and  
1289 (tv30k33pzoulql6w ?ob1)  
1290 (cql4o62p1yeke3ok )  
1291 (af9tse23ljclsqad ?ob1)  
1292 (not (a46zhwykn6jvbabk ?ob1)))  
1293 )  
1294 (:action j4gv801gnu2it0yj  
1295 :parameters (?ob1 - object ?ob2 - object)  
1296 :precondition (and (tv30k33pzoulql6w ?ob2) (a46zhwykn6jvbabk ?ob1))  
1297 :effect (and  
1298 (cql4o62p1yeke3ok )  
1299 (tv30k33pzoulql6w ?ob1)



```

1296         (b6e9q4r60gagvdcn ?ob1 ?ob2)
1297         (not (tv30k33pzoulql6w ?ob2))
1298         (not (a46zhwykn6jvbabk ?ob1)))
1299     )
1300
1301     (:action wio5amhq7814n006
1302      :parameters (?ob1 - object ?ob2 - object)
1303      :precondition (and (b6e9q4r60gagvdcn ?ob1 ?ob2) (tv30k33pzoulql6w ?ob1) (
1304        cql4o62p1yeke3ok ))
1305      :effect (and
1306        (a46zhwykn6jvbabk ?ob1)
1307        (tv30k33pzoulql6w ?ob2)
1308        (not (b6e9q4r60gagvdcn ?ob1 ?ob2))
1309        (not (tv30k33pzoulql6w ?ob1))
1310        (not (cql4o62p1yeke3ok )))
1311    )
1312
1313 [QUERY PROBLEM]
1314 (define(problem instance-1-4)
1315 (:domain xaji0y)
1316 (:objects a b c d )
1317 (:init
1318 (cql4o62p1yeke3ok)
1319 (b6e9q4r60gagvdcn a d)
1320 (af9tse23ljclsqad b)
1321 (b6e9q4r60gagvdcn c a)
1322 (b6e9q4r60gagvdcn d b)
1323 (tv30k33pzoulql6w c)
1324 )
1325 (:goal
1326 (and
1327   (b6e9q4r60gagvdcn a d)
1328   (b6e9q4r60gagvdcn b a)
1329   (b6e9q4r60gagvdcn b d)
1330 )
1331 )
1332 )
1333
1334 [PLAN]
1335 =====o1-preview RESPONSE=====
1336 There is no sequence of actions that, when run from the specified initial state,
1337 satisfies the specified goal. Thus, no plan can be provided.
1338
1339 K PROMPTS FOR LOGISTICS
1340
1341 K.1 LOGISTICS - ZERO-SHOT IN PDDL
1342
1343 The following is a PDDL specification of a planning problem. The first part, under
1344 the heading [DOMAIN], is the domain file. The second part, under the heading [
1345 QUERY PROBLEM], is the problem file. Using this information, which is correct,
1346 and no further assumptions, find a plan which, when run from the specified
1347 initial state, satisfies the specified goal. Provide your answer as a sequence
1348 of actions in PDDL format. An action ACTION which acts on two objects OBJ1 and
1349 OBJ2 would be written (ACTION OBJ1 OBJ2). Do not provide anything else in your
1350 answer.
1351 [DOMAIN]
1352 (define (domain logistics-strips)
1353 (:requirements :strips)
1354 (:predicates (OBJ ?obj)
1355              (TRUCK ?truck)
1356              (LOCATION ?loc)

```

---

```

1350             (AIRPLANE ?airplane)
1351             (CITY ?city)
1352             (AIRPORT ?airport)
1353             (at ?obj ?loc)
1354             (in ?obj1 ?obj2)
1355             (in-city ?obj ?city))
1356 ; (:types )           ; default object
1357
1358 (:action LOAD-TRUCK
1359 :parameters
1360   (?obj
1361    ?truck
1362    ?loc)
1363 :precondition
1364   (and (OBJ ?obj) (TRUCK ?truck) (LOCATION ?loc)
1365        (at ?truck ?loc) (at ?obj ?loc))
1366 :effect
1367   (and (not (at ?obj ?loc)) (in ?obj ?truck)))
1368
1369 (:action LOAD-AIRPLANE
1370 :parameters
1371   (?obj
1372    ?airplane
1373    ?loc)
1374 :precondition
1375   (and (OBJ ?obj) (AIRPLANE ?airplane) (LOCATION ?loc)
1376        (at ?obj ?loc) (at ?airplane ?loc))
1377 :effect
1378   (and (not (at ?obj ?loc)) (in ?obj ?airplane)))
1379
1380 (:action UNLOAD-TRUCK
1381 :parameters
1382   (?obj
1383    ?truck
1384    ?loc)
1385 :precondition
1386   (and (OBJ ?obj) (TRUCK ?truck) (LOCATION ?loc)
1387        (at ?truck ?loc) (in ?obj ?truck))
1388 :effect
1389   (and (not (in ?obj ?truck)) (at ?obj ?loc)))
1390
1391 (:action UNLOAD-AIRPLANE
1392 :parameters
1393   (?obj
1394    ?airplane
1395    ?loc)
1396 :precondition
1397   (and (OBJ ?obj) (AIRPLANE ?airplane) (LOCATION ?loc)
1398        (in ?obj ?airplane) (at ?airplane ?loc))
1399 :effect
1400   (and (not (in ?obj ?airplane)) (at ?obj ?loc)))
1401
1402 (:action DRIVE-TRUCK
1403 :parameters
1404   (?truck
1405    ?loc-from
1406    ?loc-to
1407    ?city)
1408 :precondition
1409   (and (TRUCK ?truck) (LOCATION ?loc-from) (LOCATION ?loc-to) (CITY ?city)
1410        (at ?truck ?loc-from)
1411        (in-city ?loc-from ?city))

```

```

1404     (in-city ?loc-to ?city))
1405   :effect
1406     (and (not (at ?truck ?loc-from)) (at ?truck ?loc-to)))
1407
1408   (:action FLY-AIRPLANE
1409     :parameters
1410       (?airplane
1411         ?loc-from
1412         ?loc-to)
1413     :precondition
1414       (and (AIRPLANE ?airplane) (AIRPORT ?loc-from) (AIRPORT ?loc-to)
1415         (at ?airplane ?loc-from))
1416     :effect
1417       (and (not (at ?airplane ?loc-from)) (at ?airplane ?loc-to)))
1418   )
1419
1420 [QUERY PROBLEM]
1421 (define(problem logistics-c2-s1-p1-a2)
1422   (:domain logistics-strips)
1423   (:objects a0 a1
1424             c0 c1
1425             t0 t1
1426             l0-0 l1-0
1427             p0)
1428   (:init
1429     (AIRPLANE a0)
1430     (AIRPLANE a1)
1431     (CITY c0)
1432     (CITY c1)
1433     (TRUCK t0)
1434     (TRUCK t1)
1435     (LOCATION l0-0)
1436     (in-city l0-0 c0)
1437     (LOCATION l1-0)
1438     (in-city l1-0 c1)
1439     (AIRPORT l0-0)
1440     (AIRPORT l1-0)
1441     (OBJ p0)
1442     (at t0 l0-0)
1443     (at t1 l1-0)
1444     (at p0 l1-0)
1445     (at a0 l0-0)
1446     (at a1 l0-0)
1447   )
1448   (:goal
1449     (and
1450       (at p0 l0-0)
1451     )
1452   )
1453 )
1454 )
1455
1456 [PLAN]
1457 =====o1-preview RESPONSE=====
1458 (FLY-AIRPLANE a0 l0-0 l1-0)
1459
1460 (LOAD-AIRPLANE p0 a0 l1-0)
1461
1462 (FLY-AIRPLANE a0 l1-0 l0-0)
1463
1464 (UNLOAD-AIRPLANE p0 a0 l0-0)

```

K.2 RANDOMIZED LOGISTICS - ZERO-SHOT IN PDDL

---

1458 The following is a PDDL specification of a planning problem. The first part,  
1459 under the heading [DOMAIN], is the domain file. The second part, under the  
1460 heading [QUERY PROBLEM], is the problem file. Using this information, which  
1461 is correct, and no further assumptions, find a plan which, when run from the  
1462 specified initial state, satisfies the specified goal. Provide your answer  
1463 as a sequence of actions in PDDL format. An action ACTION which acts on two  
1464 objects OBJ1 and OBJ2 would be written (ACTION OBJ1 OBJ2). Do not provide  
1465 anything else in your answer.

1465 [DOMAIN]  
1466 (define (domain dxpflc3up53obvi)  
1467 (:requirements :strips)  
1468 (:predicates (a46zhwykn6jvbabk ?airport) (abwaom5mhyf979wg ?obj1 ?obj2) (  
1469 b6e9q4r60gagvdcn ?city) (cql4o62p1yeke3ok ?airplane) (dvr5qjofrygv9zlx ?obj  
1470 ?loc) (fe8cgp3y92p61b85 ?loc) (j3bxzst71100cyne ?obj ?city) (  
1471 lwiiiuk750o3u28e ?obj) (rvgqqfhogwl9yspo ?truck))  
1472 (:action af9tse23ljclsqad  
1473 :parameters (?ob\_1 ?ob\_2 ?ob\_3 ?ob\_4)  
1474 :precondition (and (rvgqqfhogwl9yspo ?ob\_1) (fe8cgp3y92p61b85 ?ob\_2) (  
1475 fe8cgp3y92p61b85 ?ob\_3) (b6e9q4r60gagvdcn ?ob\_4) (dvr5qjofrygv9zlx ?ob\_1  
1476 ?ob\_2) (j3bxzst71100cyne ?ob\_2 ?ob\_4) (j3bxzst71100cyne ?ob\_3 ?ob\_4))  
1477 :effect (and (not (dvr5qjofrygv9zlx ?ob\_1 ?ob\_2)) (dvr5qjofrygv9zlx ?ob\_1 ?  
1478 ob\_3))  
1479 )  
1480 (:action j4gv801gnu2it0yj  
1481 :parameters (?ob\_1 ?ob\_2 ?ob\_3)  
1482 :precondition (and (lwiiiuk750o3u28e ?ob\_1) (rvgqqfhogwl9yspo ?ob\_2) (  
1483 fe8cgp3y92p61b85 ?ob\_3) (dvr5qjofrygv9zlx ?ob\_2 ?ob\_3) (abwaom5mhyf979wg  
1484 ?ob\_1 ?ob\_2))  
1485 :effect (and (not (abwaom5mhyf979wg ?ob\_1 ?ob\_2)) (dvr5qjofrygv9zlx ?ob\_1 ?  
1486 ob\_3))  
1487 )  
1488 (:action kip9uw781pv62umn  
1489 :parameters (?ob\_1 ?ob\_2 ?ob\_3)  
1490 :precondition (and (lwiiiuk750o3u28e ?ob\_1) (cql4o62p1yeke3ok ?ob\_2) (  
1491 fe8cgp3y92p61b85 ?ob\_3) (dvr5qjofrygv9zlx ?ob\_1 ?ob\_3) (dvr5qjofrygv9zlx  
1492 ?ob\_2 ?ob\_3))  
1493 :effect (and (not (dvr5qjofrygv9zlx ?ob\_1 ?ob\_3)) (abwaom5mhyf979wg ?ob\_1 ?  
1494 ob\_2))  
1495 )  
1496 (:action tv30k33pzoulql6w  
1497 :parameters (?ob\_1 ?ob\_2 ?ob\_3)  
1498 :precondition (and (cql4o62p1yeke3ok ?ob\_1) (a46zhwykn6jvbabk ?ob\_2) (  
1499 a46zhwykn6jvbabk ?ob\_3) (dvr5qjofrygv9zlx ?ob\_1 ?ob\_2))  
1500 :effect (and (not (dvr5qjofrygv9zlx ?ob\_1 ?ob\_2)) (dvr5qjofrygv9zlx ?ob\_1 ?  
1501 ob\_3))  
1502 )  
1503 (:action u64y1a9apusmslxb  
1504 :parameters (?ob\_1 ?ob\_2 ?ob\_3)  
1505 :precondition (and (lwiiiuk750o3u28e ?ob\_1) (rvgqqfhogwl9yspo ?ob\_2) (  
1506 fe8cgp3y92p61b85 ?ob\_3) (dvr5qjofrygv9zlx ?ob\_2 ?ob\_3) (dvr5qjofrygv9zlx  
1507 ?ob\_1 ?ob\_3))  
1508 :effect (and (not (dvr5qjofrygv9zlx ?ob\_1 ?ob\_3)) (abwaom5mhyf979wg ?ob\_1 ?  
1509 ob\_2))  
1510 )  
1511 (:action wio5amhq7814n006  
1512 :parameters (?ob\_1 ?ob\_2 ?ob\_3)  
1513 :precondition (and (lwiiiuk750o3u28e ?ob\_1) (cql4o62p1yeke3ok ?ob\_2) (  
1514 fe8cgp3y92p61b85 ?ob\_3) (abwaom5mhyf979wg ?ob\_1 ?ob\_2) (dvr5qjofrygv9zlx  
1515 ?ob\_2 ?ob\_3))  
1516 :effect (and (not (abwaom5mhyf979wg ?ob\_1 ?ob\_2)) (dvr5qjofrygv9zlx ?ob\_1 ?  
1517 ob\_3))  
1518 )  
1519 )  
1520 )  
1521 )

[QUERY PROBLEM]

```

1512 (define(problem vd24g3dfuion115)
1513   (:domain dxpflc3up53obvi)
1514   (:objects o1 o2 o3 o4 o5 o6 o7 o8 o9)
1515   (:init (a46zhwykn6jvbabk o4) (a46zhwykn6jvbabk o7) (b6e9q4r60gagvdcn o1) (
1516     b6e9q4r60gagvdcn o9) (cql4o62plyeke3ok o2) (cql4o62plyeke3ok o6) (
1517     dvr5qjofrygv9zlx o2 o7) (dvr5qjofrygv9zlx o3 o4) (dvr5qjofrygv9zlx o5 o4) (
1518     dvr5qjofrygv9zlx o6 o7) (dvr5qjofrygv9zlx o8 o7) (fe8cgp3y92p61b85 o4) (
1519     fe8cgp3y92p61b85 o7) (j3bxzst71100cyne o4 o9) (j3bxzst71100cyne o7 o1) (
1520     lwiiiuk750o3u28e o5) (rvgqqfhogwl9yspo o3) (rvgqqfhogwl9yspo o8))
1521   (:goal (dvr5qjofrygv9zlx o5 o7))
1522 )
1523 [PLAN]
1524 =====o1-preview RESPONSE=====
1525 (tv30k33pzoulql6w o2 o7 o4)
1526 (kip9uw781pv62umn o5 o2 o4)
1527 (tv30k33pzoulql6w o2 o4 o7)
1528 (wio5amhq7814n006 o5 o2 o7)
1529
1530 L PROMPTS FOR SOKOBAN
1531
1532 L.1 SOKOBAN - ZERO-SHOT IN PDDL
1533
1534 The following is a PDDL specification of a planning problem. The first part, under
1535 the heading [DOMAIN], is the domain file. The second part, under the heading [
1536 QUERY PROBLEM], is the problem file. Using this information, which is correct,
1537 and no further assumptions, find a plan which, when run from the specified
1538 initial state, satisfies the specified goal. Provide your answer as a sequence
1539 of actions in PDDL format. An action ACTION which acts on two objects OBJ1 and
1540 OBJ2 would be written (ACTION OBJ1 OBJ2). Do not provide anything else in your
1541 answer.
1542 [DOMAIN]
1543 (define (domain typed-sokoban)
1544   (:requirements :typing)
1545   (:types LOC DIR BOX)
1546   (:predicates
1547     (at-robot ?l - LOC)
1548     (at ?o - BOX ?l - LOC)
1549     (adjacent ?l1 - LOC ?l2 - LOC ?d - DIR)
1550     (clear ?l - LOC)
1551   )
1552   (:action move
1553     :parameters (?from - LOC ?to - LOC ?dir - DIR)
1554     :precondition (and (clear ?to) (at-robot ?from) (adjacent ?from ?to ?dir))
1555     :effect (and (at-robot ?to) (not (at-robot ?from)))
1556   )
1557   (:action push
1558     :parameters (?rloc - LOC ?bloc - LOC ?floc - LOC ?dir - DIR ?b - BOX)
1559     :precondition (and (at-robot ?rloc) (at ?b ?bloc) (clear ?floc)
1560       (adjacent ?rloc ?bloc ?dir) (adjacent ?bloc ?floc ?dir))
1561     :effect (and (at-robot ?bloc) (at ?b ?floc) (clear ?bloc)
1562       (not (at-robot ?rloc)) (not (at ?b ?bloc)) (not (clear ?floc)))
1563   )
1564 )
1565 [QUERY PROBLEM]
1566 (define(problem typed-sokoban-grid7-boxes1-walls2)
1567   (:domain typed-sokoban)
1568   (:objects

```

```

1566         up down left right - DIR
1567         box0 - BOX
1568         f0-0f f0-1f f0-2f f0-3f f0-4f f0-5f f0-6f
1569         f1-0f f1-1f f1-2f f1-3f f1-4f f1-5f f1-6f
1570         f2-0f f2-1f f2-2f f2-3f f2-4f f2-5f f2-6f
1571         f3-0f f3-1f f3-2f f3-3f f3-4f f3-5f f3-6f
1572         f4-0f f4-1f f4-2f f4-3f f4-4f f4-5f f4-6f
1573         f5-0f f5-1f f5-2f f5-3f f5-4f f5-5f f5-6f
1574         f6-0f f6-1f f6-2f f6-3f f6-4f f6-5f f6-6f - LOC
1574     )
1575     (:init
1576     (adjacent f0-0f f0-1f right)
1577     (adjacent f0-0f f1-0f down)
1578     (adjacent f0-1f f0-0f left)
1579     (adjacent f0-1f f0-2f right)
1580     (adjacent f0-1f f1-1f down)
1581     (adjacent f0-2f f0-1f left)
1582     (adjacent f0-2f f0-3f right)
1583     (adjacent f0-2f f1-2f down)
1584     (adjacent f0-3f f0-2f left)
1585     (adjacent f0-3f f0-4f right)
1586     (adjacent f0-3f f1-3f down)
1587     (adjacent f0-4f f0-3f left)
1588     (adjacent f0-4f f0-5f right)
1589     (adjacent f0-4f f1-4f down)
1590     (adjacent f0-5f f0-4f left)
1591     (adjacent f0-5f f0-6f right)
1592     (adjacent f0-5f f1-5f down)
1593     (adjacent f0-6f f0-5f left)
1594     (adjacent f0-6f f1-6f down)
1595     (adjacent f1-0f f1-1f right)
1596     (adjacent f1-0f f0-0f up)
1597     (adjacent f1-0f f2-0f down)
1598     (adjacent f1-1f f1-0f left)
1599     (adjacent f1-1f f1-2f right)
1600     (adjacent f1-1f f0-1f up)
1601     (adjacent f1-1f f2-1f down)
1602     (adjacent f1-2f f1-1f left)
1603     (adjacent f1-2f f1-3f right)
1604     (adjacent f1-2f f0-2f up)
1605     (adjacent f1-2f f2-2f down)
1606     (adjacent f1-3f f1-2f left)
1607     (adjacent f1-3f f1-4f right)
1608     (adjacent f1-3f f0-3f up)
1609     (adjacent f1-3f f2-3f down)
1610     (adjacent f1-4f f1-3f left)
1611     (adjacent f1-4f f1-5f right)
1612     (adjacent f1-4f f0-4f up)
1613     (adjacent f1-4f f2-4f down)
1614     (adjacent f1-5f f1-4f left)
1615     (adjacent f1-5f f1-6f right)
1616     (adjacent f1-5f f0-5f up)
1617     (adjacent f1-5f f2-5f down)
1618     (adjacent f1-6f f1-5f left)
1619     (adjacent f1-6f f0-6f up)
1620     (adjacent f1-6f f2-6f down)
1621     (adjacent f2-0f f2-1f right)
1622     (adjacent f2-0f f1-0f up)
1623     (adjacent f2-0f f3-0f down)
1624     (adjacent f2-1f f2-0f left)
1625     (adjacent f2-1f f2-2f right)
1626     (adjacent f2-1f f1-1f up)
1627     (adjacent f2-1f f3-1f down)
1628     (adjacent f2-2f f2-1f left)
1629     (adjacent f2-2f f2-3f right)

```

---

1620 (adjacent f2-2f f1-2f up)  
1621 (adjacent f2-2f f3-2f down)  
1622 (adjacent f2-3f f2-2f left)  
1623 (adjacent f2-3f f2-4f right)  
1624 (adjacent f2-3f f1-3f up)  
1625 (adjacent f2-3f f3-3f down)  
1626 (adjacent f2-4f f2-3f left)  
1627 (adjacent f2-4f f2-5f right)  
1628 (adjacent f2-4f f1-4f up)  
1629 (adjacent f2-4f f3-4f down)  
1630 (adjacent f2-5f f2-4f left)  
1631 (adjacent f2-5f f2-6f right)  
1632 (adjacent f2-5f f1-5f up)  
1633 (adjacent f2-5f f3-5f down)  
1634 (adjacent f2-6f f2-5f left)  
1635 (adjacent f2-6f f1-6f up)  
1636 (adjacent f2-6f f3-6f down)  
1637 (adjacent f3-0f f3-1f right)  
1638 (adjacent f3-0f f2-0f up)  
1639 (adjacent f3-0f f4-0f down)  
1640 (adjacent f3-1f f3-0f left)  
1641 (adjacent f3-1f f3-2f right)  
1642 (adjacent f3-1f f2-1f up)  
1643 (adjacent f3-1f f4-1f down)  
1644 (adjacent f3-2f f3-1f left)  
1645 (adjacent f3-2f f3-3f right)  
1646 (adjacent f3-2f f2-2f up)  
1647 (adjacent f3-2f f4-2f down)  
1648 (adjacent f3-3f f3-2f left)  
1649 (adjacent f3-3f f3-4f right)  
1650 (adjacent f3-3f f2-3f up)  
1651 (adjacent f3-3f f4-3f down)  
1652 (adjacent f3-4f f3-3f left)  
1653 (adjacent f3-4f f3-5f right)  
1654 (adjacent f3-4f f2-4f up)  
1655 (adjacent f3-4f f4-4f down)  
1656 (adjacent f3-5f f3-4f left)  
1657 (adjacent f3-5f f3-6f right)  
1658 (adjacent f3-5f f2-5f up)  
1659 (adjacent f3-5f f4-5f down)  
1660 (adjacent f3-6f f3-5f left)  
1661 (adjacent f3-6f f2-6f up)  
1662 (adjacent f3-6f f4-6f down)  
1663 (adjacent f4-0f f4-1f right)  
1664 (adjacent f4-0f f3-0f up)  
1665 (adjacent f4-0f f5-0f down)  
1666 (adjacent f4-1f f4-0f left)  
1667 (adjacent f4-1f f4-2f right)  
1668 (adjacent f4-1f f3-1f up)  
1669 (adjacent f4-1f f5-1f down)  
1670 (adjacent f4-2f f4-1f left)  
1671 (adjacent f4-2f f4-3f right)  
1672 (adjacent f4-2f f3-2f up)  
1673 (adjacent f4-2f f5-2f down)  
1674 (adjacent f4-3f f4-2f left)  
1675 (adjacent f4-3f f4-4f right)  
1676 (adjacent f4-3f f3-3f up)  
1677 (adjacent f4-3f f5-3f down)  
1678 (adjacent f4-4f f4-3f left)  
1679 (adjacent f4-4f f4-5f right)  
1680 (adjacent f4-4f f3-4f up)  
1681 (adjacent f4-4f f5-4f down)  
1682 (adjacent f4-5f f4-4f left)  
1683 (adjacent f4-5f f4-6f right)  
1684 (adjacent f4-5f f3-5f up)

---

1674 (adjacent f4-5f f5-5f down)  
1675 (adjacent f4-6f f4-5f left)  
1676 (adjacent f4-6f f3-6f up)  
1677 (adjacent f4-6f f5-6f down)  
1678 (adjacent f5-0f f5-1f right)  
1679 (adjacent f5-0f f4-0f up)  
1680 (adjacent f5-0f f6-0f down)  
1681 (adjacent f5-1f f5-0f left)  
1682 (adjacent f5-1f f5-2f right)  
1683 (adjacent f5-1f f4-1f up)  
1684 (adjacent f5-1f f6-1f down)  
1685 (adjacent f5-2f f5-1f left)  
1686 (adjacent f5-2f f5-3f right)  
1687 (adjacent f5-2f f4-2f up)  
1688 (adjacent f5-2f f6-2f down)  
1689 (adjacent f5-3f f5-2f left)  
1690 (adjacent f5-3f f5-4f right)  
1691 (adjacent f5-3f f4-3f up)  
1692 (adjacent f5-3f f6-3f down)  
1693 (adjacent f5-4f f5-3f left)  
1694 (adjacent f5-4f f5-5f right)  
1695 (adjacent f5-4f f4-4f up)  
1696 (adjacent f5-4f f6-4f down)  
1697 (adjacent f5-5f f5-4f left)  
1698 (adjacent f5-5f f5-6f right)  
1699 (adjacent f5-5f f4-5f up)  
1700 (adjacent f5-5f f6-5f down)  
1701 (adjacent f5-6f f5-5f left)  
1702 (adjacent f5-6f f4-6f up)  
1703 (adjacent f5-6f f6-6f down)  
1704 (adjacent f6-0f f6-1f right)  
1705 (adjacent f6-0f f5-0f up)  
1706 (adjacent f6-1f f6-0f left)  
1707 (adjacent f6-1f f6-2f right)  
1708 (adjacent f6-1f f5-1f up)  
1709 (adjacent f6-2f f6-1f left)  
1710 (adjacent f6-2f f6-3f right)  
1711 (adjacent f6-2f f5-2f up)  
1712 (adjacent f6-3f f6-2f left)  
1713 (adjacent f6-3f f6-4f right)  
1714 (adjacent f6-3f f5-3f up)  
1715 (adjacent f6-4f f6-3f left)  
1716 (adjacent f6-4f f6-5f right)  
1717 (adjacent f6-4f f5-4f up)  
1718 (adjacent f6-5f f6-4f left)  
1719 (adjacent f6-5f f6-6f right)  
1720 (adjacent f6-5f f5-5f up)  
1721 (adjacent f6-6f f6-5f left)  
1722 (adjacent f6-6f f5-6f up)  
1723 (at box0 f1-2f)  
1724 (clear f0-0f)  
1725 (clear f0-2f)  
1726 (clear f0-3f)  
1727 (clear f0-4f)  
1728 (clear f0-5f)  
1729 (clear f0-6f)  
1730 (clear f1-0f)  
1731 (clear f1-1f)  
1732 (clear f1-3f)  
1733 (clear f1-4f)  
1734 (clear f1-5f)  
1735 (clear f1-6f)  
1736 (clear f2-0f)  
1737 (clear f2-1f)  
1738 (clear f2-2f)



```

1728 (clear f2-3f)
1729 (clear f2-4f)
1730 (clear f2-5f)
1731 (clear f2-6f)
1732 (clear f3-0f)
1733 (clear f3-3f)
1734 (clear f3-4f)
1735 (clear f3-5f)
1736 (clear f3-6f)
1737 (clear f4-1f)
1738 (clear f4-2f)
1739 (clear f4-3f)
1740 (clear f4-4f)
1741 (clear f4-5f)
1742 (clear f4-6f)
1743 (clear f5-0f)
1744 (clear f5-1f)
1745 (clear f5-2f)
1746 (clear f5-3f)
1747 (clear f5-4f)
1748 (clear f5-5f)
1749 (clear f5-6f)
1750 (at-robot f6-0f)
1751 (clear f6-0f)
1752 (clear f6-1f)
1753 (clear f6-2f)
1754 (clear f6-3f)
1755 (clear f6-4f)
1756 (clear f6-6f)
1757 )
1758 (:goal
1759 (and
1760 (at box0 f6-6f)
1761 )
1762 )
1763 [PLAN]
1764 =====o1-preview RESPONSE=====
1765 (move f6-0f f5-0f up)
1766 (move f5-0f f4-0f up)
1767 (move f4-0f f3-0f up)
1768 (move f3-0f f2-0f up)
1769 (move f2-0f f1-0f up)
1770 (move f1-0f f1-1f right)
1771 (push f1-1f f1-2f f1-3f right box0)
1772 (push f1-2f f1-3f f1-4f right box0)
1773 (push f1-3f f1-4f f1-5f right box0)
1774 (push f1-4f f1-5f f1-6f right box0)
1775 (move f1-5f f0-5f up)
1776 (move f0-5f f0-6f right)
1777 (push f0-6f f1-6f f2-6f down box0)
1778 (push f1-6f f2-6f f3-6f down box0)
1779 (push f2-6f f3-6f f4-6f down box0)
1780 (push f3-6f f4-6f f5-6f down box0)
1781 (push f4-6f f5-6f f6-6f down box0)
1782
1783 L.2 SOKOBAN - BACKPROMPT
1784
1785 The above plan is invalid. The following action at step 20 has an unmet precondition
1786 :
1787 push f8-5f f8-6f f8-7f right box1
1788 The unmet precondition is:

```

---

1782 (clear f8-7f)  
1783 Provide your revised plan as a sequence of actions in PDDL format. An action ACTION  
1784 which acts on two objects OBJ1 and OBJ2 would be written (ACTION OBJ1 OBJ2). Do  
1785 not provide anything else in your answer.  
1786  
1787 [PLAN]  
1788  
1789 M PROMPTS FOR GRAPH COLORING  
1790  
1791 M.1 GRAPH COLORING - HARD  
1792  
1793 Color the following graph, described as a set of edges, such that no two  
1794 vertices on the same edge share a color.  
1795 You may use at most 5 colors.  
1796 Vertex 0 is connected to vertex 3.  
1797 Vertex 0 is connected to vertex 6.  
1798 Vertex 0 is connected to vertex 8.  
1799 Vertex 0 is connected to vertex 13.  
1800 Vertex 0 is connected to vertex 15.  
1801 Vertex 0 is connected to vertex 17.  
1802 Vertex 0 is connected to vertex 19.  
1803 Vertex 1 is connected to vertex 4.  
1804 Vertex 1 is connected to vertex 5.  
1805 Vertex 1 is connected to vertex 8.  
1806 Vertex 1 is connected to vertex 11.  
1807 Vertex 1 is connected to vertex 12.  
1808 Vertex 1 is connected to vertex 13.  
1809 Vertex 1 is connected to vertex 14.  
1810 Vertex 1 is connected to vertex 17.  
1811 Vertex 1 is connected to vertex 18.  
1812 Vertex 1 is connected to vertex 19.  
1813 Vertex 2 is connected to vertex 3.  
1814 Vertex 2 is connected to vertex 4.  
1815 Vertex 2 is connected to vertex 6.  
1816 Vertex 2 is connected to vertex 8.  
1817 Vertex 2 is connected to vertex 13.  
1818 Vertex 2 is connected to vertex 14.  
1819 Vertex 2 is connected to vertex 15.  
1820 Vertex 2 is connected to vertex 16.  
1821 Vertex 2 is connected to vertex 17.  
1822 Vertex 3 is connected to vertex 5.  
1823 Vertex 3 is connected to vertex 8.  
1824 Vertex 3 is connected to vertex 11.  
1825 Vertex 3 is connected to vertex 12.  
1826 Vertex 3 is connected to vertex 14.  
1827 Vertex 3 is connected to vertex 16.  
1828 Vertex 4 is connected to vertex 5.  
1829 Vertex 4 is connected to vertex 7.  
1830 Vertex 4 is connected to vertex 9.  
1831 Vertex 4 is connected to vertex 11.  
1832 Vertex 4 is connected to vertex 17.  
1833 Vertex 4 is connected to vertex 18.  
1834 Vertex 5 is connected to vertex 11.  
1835 Vertex 5 is connected to vertex 14.  
1836 Vertex 5 is connected to vertex 15.  
1837 Vertex 6 is connected to vertex 11.  
1838 Vertex 6 is connected to vertex 16.  
1839 Vertex 6 is connected to vertex 17.  
1840 Vertex 7 is connected to vertex 9.  
1841 Vertex 7 is connected to vertex 10.  
1842 Vertex 7 is connected to vertex 13.  
1843 Vertex 7 is connected to vertex 14.  
1844 Vertex 7 is connected to vertex 16.  
1845 Vertex 7 is connected to vertex 17.

---

1836 Vertex 8 is connected to vertex 10.  
1837 Vertex 8 is connected to vertex 12.  
1838 Vertex 8 is connected to vertex 13.  
1839 Vertex 8 is connected to vertex 16.  
1840 Vertex 8 is connected to vertex 19.  
1841 Vertex 9 is connected to vertex 11.  
1842 Vertex 9 is connected to vertex 13.  
1843 Vertex 9 is connected to vertex 17.  
1844 Vertex 10 is connected to vertex 11.  
1845 Vertex 10 is connected to vertex 13.  
1846 Vertex 10 is connected to vertex 16.  
1847 Vertex 10 is connected to vertex 17.  
1848 Vertex 10 is connected to vertex 18.  
1849 Vertex 10 is connected to vertex 19.  
1850 Vertex 11 is connected to vertex 13.  
1851 Vertex 11 is connected to vertex 17.  
1852 Vertex 11 is connected to vertex 18.  
1853 Vertex 11 is connected to vertex 19.  
1854 Vertex 12 is connected to vertex 13.  
1855 Vertex 12 is connected to vertex 15.  
1856 Vertex 12 is connected to vertex 16.  
1857 Vertex 12 is connected to vertex 18.  
1858 Vertex 12 is connected to vertex 19.  
1859 Vertex 13 is connected to vertex 14.  
1860 Vertex 13 is connected to vertex 18.  
1861 Vertex 14 is connected to vertex 15.  
1862 Vertex 14 is connected to vertex 18.  
1863 Vertex 14 is connected to vertex 19.  
1864 There are a total of 20 vertices. Please label every vertex, even if it is  
1865 disconnected from the rest of the graph. Please provide each vertex's color. Do  
1866 not skip any vertices. Each color must be provided on a new line in the  
1867 response and should be formatted as "{VERTEX NUMBER}: {VERTEX COLOR ASSIGNMENT  
1868 }". Please do not provide anything else in your response, and end your response  
1869 with '[ANSWER END]'

1870 =====o1-preview RESPONSE=====

1871 0: 1  
1872 1: 3  
1873 2: 1  
1874 3: 2  
1875 4: 4  
1876 5: 5  
1877 6: 3  
1878 7: 1  
1879 8: 4  
1880 9: 3  
1881 10: 3  
1882 11: 1  
1883 12: 1  
1884 13: 2  
1885 14: 4  
1886 15: 2  
1887 16: 5  
1888 17: 2  
1889 18: 5  
1890 19: 2  
1891 [ANSWER END]

1892

1893 N GRAPH COLORING BACKPROMPT - HARD

1894

1895 This is incorrect. Feedback:  
1896 Vertex 0 was not given a value in the coloring.  
1897 Vertex 1 was not given a value in the coloring.  
1898 Vertex 0 was not given a value in the coloring.  
1899 Vertex 2 was not given a value in the coloring.





---

1998 Vertex 9 was not given a value in the coloring.  
1999 Vertex 15 was not given a value in the coloring.  
2000 Vertex 10 was not given a value in the coloring.  
2001 Vertex 12 was not given a value in the coloring.  
2002 Vertex 10 was not given a value in the coloring.  
2003 Vertex 16 was not given a value in the coloring.  
2004 Vertex 10 was not given a value in the coloring.  
2005 Vertex 17 was not given a value in the coloring.  
2006 Vertex 11 was not given a value in the coloring.  
2007 Vertex 12 was not given a value in the coloring.  
2008 Vertex 11 was not given a value in the coloring.  
2009 Vertex 14 was not given a value in the coloring.  
2010 Vertex 11 was not given a value in the coloring.  
2011 Vertex 15 was not given a value in the coloring.  
2012 Vertex 11 was not given a value in the coloring.  
2013 Vertex 17 was not given a value in the coloring.  
2014 Vertex 11 was not given a value in the coloring.  
2015 Vertex 19 was not given a value in the coloring.  
2016 Vertex 12 was not given a value in the coloring.  
2017 Vertex 13 was not given a value in the coloring.  
2018 Vertex 12 was not given a value in the coloring.  
2019 Vertex 18 was not given a value in the coloring.  
2020 Vertex 13 was not given a value in the coloring.  
2021 Vertex 17 was not given a value in the coloring.  
2022 Vertex 14 was not given a value in the coloring.  
2023 Vertex 19 was not given a value in the coloring.  
2024 Vertex 16 was not given a value in the coloring.  
2025 Vertex 17 was not given a value in the coloring.  
2026 Vertex 18 was not given a value in the coloring.  
2027  
2028 Using this feedback, please try again.  
2029 Please provide each vertex's color. Do not skip any vertices. Each color must be  
2030 provided on a new line in the response and should be formatted as "{VERTEX  
2031 NUMBER}: {VERTEX COLOR ASSIGNMENT}". Please do not provide anything else in  
2032 your response, and end your response with '[ANSWER END]'  
2033  
2034 O PROMPTS FOR OSU TRAVEL PLANNING  
2035 O.1 FIRST ITERATION  
2036  
2037 You are a proficient planner. Based on the provided information and query, please  
2038 give me a detailed plan, including specifics such as flight numbers (e.g.,  
2039 F0123456),  
2040 restaurant names, and accommodation names. Note that all the information in your  
2041 plan should be derived from the provided data. You should give a travel plan in  
2042 JSON format as shown in the  
2043 example below. Additionally, all details should align with commonsense. The symbol  
2044 '-' indicates that information is unnecessary. For example, in the provided  
2045 sample, you do not need to  
2046 plan after returning to the departure city. When you travel to two cities in one day  
2047 , you should note it in the 'current\_city' section as in the example (i.e.,  
2048 from A to B).  
2049 \*\*\*\*\* Example \*\*\*\*\*  
2050 Query: Could you create a travel plan for 7 people from Ithaca to Charlotte spanning  
2051 3 days, from March 8th to March 14th, 2022, with a budget of \$30,200?  
Travel Plan:  
[  
{

```

2052     "day": 1,
2053     "people_number": 7,
2054     "current_city": "from Ithaca to Charlotte",
2055     "transportation": "Flight Number: F3633413, from Ithaca to Charlotte,
2056         Departure Time: 05:38, Arrival Time: 07:46",
2057     "breakfast": "Nagaland's Kitchen, Charlotte",
2058     "attraction": "The Charlotte Museum of History, Charlotte",
2059     "lunch": "Cafe Maple Street, Charlotte",
2060     "dinner": "Bombay Vada Pav, Charlotte",
2061     "accommodation": "Affordable Spacious Refurbished Room in Bushwick!,
2062         Charlotte"
2063 },
2064 {
2065     "day": 2,
2066     "people_number": 7,
2067     "current_city": "Charlotte",
2068     "transportation": "-",
2069     "breakfast": "Olive Tree Cafe, Charlotte",
2070     "attraction": "The Mint Museum, Charlotte;Romare Bearden Park, Charlotte",
2071     "lunch": "Birbal Ji Dhaba, Charlotte",
2072     "dinner": "Pind Balluchi, Charlotte",
2073     "accommodation": "Affordable Spacious Refurbished Room in Bushwick!,
2074         Charlotte"
2075 },
2076 {
2077     "day": 3,
2078     "people_number": 7,
2079     "current_city": "from Charlotte to Ithaca",
2080     "transportation": "Flight Number: F3786167, from Charlotte to Ithaca,
2081         Departure Time: 21:42, Arrival Time: 23:26",
2082     "breakfast": "Books Monument, Charlotte",
2083     "attraction": "Books Monument, Charlotte",
2084     "lunch": "Olive Tree Cafe, Charlotte",
2085     "dinner": "Kylin Skybar, Charlotte",
2086     "accommodation": "-"
2087 }
2088 ]
2089 ***** Example Ends *****
2090
2091 Given information: [{'Description': 'Attractions in Nashville', 'Content': '
2092     Name Latitude Longitude
2093     Address
2094     Phone
2095     Website City\nCountry Music Hall of Fame and Museum 36.158263 -86.776126 222 Rep.
2096     John Lewis Way S, Nashville, TN 37203, USA (615) 416-2001
2097     https://countrymusichalloffame.org/ Nashville\n Nashville Zoo at Grassmere
2098     36.089705 -86.742096 3777 Nolensville Pk, Nashville, TN 37211, USA (615)
2099     833-1534
2100     http://www.nashvillezoo.org/ Nashville\n Belle Meade Historic Site & Winery
2101     36.104916 -86.864695 5025 Harding Pike, Nashville, TN 37205, USA (615) 356-0501
2102     https://visitbellemeade.com/ Nashville\n Johnny Cash Museum 36.160939
2103     -86.775757 119 3rd Ave S, Nashville, TN 37201, USA (615) 256-1777
2104     http://www.johnnycashmuseum.com/ Nashville\n Centennial Park
2105     36.148946 -86.812750 2500 West End Ave, Nashville, TN 37203, USA (615)
2106     862-8400
2107     https://www.nashville.gov/Parks-and-Recreation/Parks/Centennial-Park.aspx Nashville\n
2108     n Grand Ole Opry 36.206857 -86.692108 600 Opry Mills Dr,
2109     Nashville, TN
2110     37214, USA (615) 871-6779
2111     https://www.opry.com/?utm_campaign=opry&utm_medium=organicsearch&utm_source=
2112     googlemysbusiness&utm_audience=tofu_googlemysbusiness&utm_content=
2113     brandstory_google_my_business_website_link
2114     Nashville\n Frist Art Museum 36.157897 -86.783853 919
2115     Broadway, Nashville, TN 37203, USA (615) 244-3340

```

2106 <https://fristartmuseum.org/> Nashville\n The Parthenon 36.149674  
2107 -86.813347 2500 West End Ave, Nashville, TN 37203, USA (615) 862-8431  
2108 <https://www.nashvilleparthenon.com/> Nashville\n Nashville Shores  
2109 36.158209 -86.604915 4001 Bell Rd, Hermitage, TN 37076, USA (615)  
2110 889-7050  
2111 <http://www.nashvilleshores.com/lodging> Nashville\n Musicians Hall of Fame and Museum  
2112 36.167668 -86.782399 401 Gay St, Nashville, TN 37219, USA (615) 244-3263  
2113 <https://www.musicianshalloffame.com/> Nashville\n Bicentennial Capitol Mall State  
2114 Park 36.170887 -86.787589 600 James Robertson Pkwy, Nashville, TN 37243, USA  
2115 (888) 867-2757  
2116 <https://tnstateparks.com/parks/bicentennial-mall> Nashville\n Lane Motor  
2117 Museum 36.140197 -86.734580 702 Murfreesboro Pike, Nashville, TN 37210, USA  
2118 (615) 742-7445  
2119 <http://www.lanemotormuseum.org/> Nashville\n Honky Tonk Highway 36.160393  
2120 -86.778432 501 Broadway, Nashville, TN 37203, USA (800) 657-6910  
2121 <https://www.visitmusiccity.com/visitors/honkytonkhighway> Nashville\n Adventure  
2122 Science Center 36.146614 -86.775481 800 Fort Negley Blvd, Nashville, TN 37203,  
2123 USA (615)  
2124 862-5160  
2125 <http://www.adventuresci.org/> Nashville\n Patsy Cline Museum 36.160904  
2126 -86.776030 119 3rd Ave S, Nashville, TN 37201, USA (615) 454-4722  
2127 <https://www.patsymuseum.com/> Nashville\n Andrew Jacksons Hermitage 36.213757  
2128 -86.615310 4580 Rachels Ln, Hermitage, TN 37076, USA (615) 889-2941  
2129 <http://www.thehermitage.com/> Nashville\n NashTrash Tours 36.169474  
2130 -86.788254 900 Rosa L Parks Blvd, Nashville, TN 37208, USA (615) 226-7300  
2131 <http://www.nashtrash.com/> Nashville\n Fort Nashborough 36.164268  
2132 -86.775430 170 1st Ave N, Nashville, TN 37201, USA (615) 862-8400  
2133 <https://www.nashville.gov/Parks-and-Recreation/Historic-Sites/Fort-Nashborough.aspx>  
2134 Nashville\n Ryman Auditorium 36.161248 -86.778471 116 5th Ave N,  
2135 Nashville, TN 37219, USA (615) 889-3060  
2136 [https://www.ryman.com/?utm\\_campaign=ryman&utm\\_medium=organicsearch&utm\\_source=](https://www.ryman.com/?utm_campaign=ryman&utm_medium=organicsearch&utm_source=googlemybusiness&utm_audience=tofu_googlemybusiness&utm_content=brandstory_google_my_business_website_link)  
2137 [googlemybusiness&utm\\_audience=tofu\\_googlemybusiness&utm\\_content=](https://www.ryman.com/?utm_campaign=ryman&utm_medium=organicsearch&utm_source=googlemybusiness&utm_audience=tofu_googlemybusiness&utm_content=brandstory_google_my_business_website_link)  
2138 [brandstory\\_google\\_my\\_business\\_website\\_link](https://www.ryman.com/?utm_campaign=ryman&utm_medium=organicsearch&utm_source=googlemybusiness&utm_audience=tofu_googlemybusiness&utm_content=brandstory_google_my_business_website_link)  
2139 Nashville\n Nashville Public Square Park 36.166611 -86.778126 Union St & 3rd  
2140 Ave N, Nashville, TN 37201, USA (615) 743-3090  
2141 <https://www.nashvilledowntown.com/go/public-square-park> Nashville'}, {'Description':  
2142 'Restaurants in Nashville', 'Content': "  
2143 Cuisines Aggregate Rating City\n Bangkok 1 44  
2144 Cafe, Seafood 3.3 Nashville\n  
2145 Bablu Fast  
2146 Food 13 Desserts, Pizza, Mexican, BBQ, Fast Food 0.0 Nashville\  
2147 n Full Dabba 77 Pizza, Mexican, Fast Food, Cafe,  
2148 American  
2149 0.0 Nashville\n Twigly 14 Pizza, French, Fast Food,  
2150 Chinese, Seafood 4.5 Nashville\n Veg Hut  
2151 35  
2152 Tea, Mexican, Seafood 0.0 Nashville\n The Toddy Shop  
2153 92 Desserts, Fast Food 3.4 Nashville\  
2154 GoGourmet 34 Tea, Cafe, Indian, Mediterranean 3.7  
2155 Nashville\n Govinda's Confectionery 15 Tea, Bakery,  
2156 Mediterranean, Fast Food  
2157 3.1 Nashville\n Kitchen King 13 Pizza, Italian, BBQ,  
2158 Cafe, Mediterranean 0.0 Nashville\  
2159 55 Town Hall  
2160 Tea, Seafood 3.8 Nashville\n Smoke House Deli 59  
2161 Cafe, Mediterranean, Seafood 4.0 Nashville\  
2162 Dialogue  
2163 Lounge & Caf 83 Tea, Indian, BBQ, Cafe, American, Seafood 3.0  
2164 Nashville\n Meenakshi Bhawan 24  
2165 Tea,  
2166 Desserts 3.1 Nashville\n Oh! Calcutta 83  
2167 Tea, Mexican, BBQ, Cafe, Indian 4.4 Nashville\  
2168 Chicago  
2169



2160 Pizza 20 Tea, French, Pizza, BBQ 3.2 Nashville\  
2161 n Malhotra Restaurant 76 Tea, Desserts,  
2162 Seafood  
2163 2.9 Nashville\  
2164 Desserts, Seafood Madras Cafe 88 Pizza,  
63 0.0 Nashville\Dakshin - Sheraton New Delhi Hotel  
2165 Cafe, Pizza, Indian, Bakery 4.0 Nashville\  
2166 Imperial 76 Desserts, Pizza, Italian, American, Seafood 3.2  
2167 Nashville\  
2168 Punjabi Virsa 35 Desserts, Italian, BBQ, Cafe, Seafood 0.0  
2169 Nashville\  
2170 Cafe, Sagar Dhaba 55 Tea, Bakery, BBQ,  
2171 Mediterranean 0.0 Nashville\  
2172 Kettle & Kegs 60  
2173 Kargo 84 Tea, Pizza, Desserts 4.4 Nashville  
2174 NAME price room type house\_rules minimum nights  
2175 maximum occupancy review rate number city\nHuge 2 Bedroom, Great Location,  
2176 Express Metro 745.0 Entire home/apt No smoking & No children under 10 & No pets 4.0  
2177 4 5.0 Nashville\  
2178 private house 474.0 Clean and large bedroom in a  
2179 Private room No smoking 1.0 1  
2180 4.0 Nashville\  
2181 Entire Brooklyn Heights gem 993.0  
2182 home/apt No pets 2.0 6  
2183 Private room 5.0 Nashville\  
2184 No pets 2.0 1 4.0 Nashville\  
2185 vibrant, historic, sleeps 4 679.0 Entire home/apt No  
2186 pets  
2187 5.0 2 2.0 Nashville\  
2188 East River 870.0 Private room FiDi Cozy room overlooking  
2189 1.0 No parties  
2190 2 5.0 Nashville\  
2191 Entire home/apt No children under 10 & No pets & No smoking 3.0  
2192 2  
2193 4.0 Nashville\  
2194 Cozy bedroom close to Manhattan 576.0 Private room  
2195 No children under 10 1.0 1  
2196 3.0  
2197 Nashville'}, {'Description': 'Attractions in Knoxville', 'Content': "  
2198 Name Latitude Longitude  
2199 Address  
2200 Phone  
2201 Website City\n World's Fair Park 35.962577 -83.924192  
2202 525 Henley St, Knoxville, TN 37902, USA (865) 215-1158  
2203 http:// worldsfairpark.org/ Knoxville\  
2204 Knoxville Museum of Art 35.962426 -83.925229 1050 Worlds Fair Park Dr, Knoxville, TN  
2205 37916, USA (865) 525-6101  
2206 http://www.knoxart.org/ Knoxville\  
2207 -83.923353 810 Clinch Ave, Knoxville, TN 37902, USA (865) 314-0660  
2208 http://www.sunspheretickets.com/ Knoxville\  
2209 35.956454 -83.866775 2915 Island Home Ave, Knoxville, TN 37920, USA (865)  
2210 577-4717  
2211 http://www.ijams.org/ Knoxville\  
2212 -83.919167 301 S Gay St, Knoxville, TN 37902, USA (865) 309-4522  
2213 http://knoxvillewalkingtours.com/ Knoxville\  
2214 35.997617 -83.885467 516 N Beaman St, Knoxville, TN 37914, USA (865)  
2215 594-1494  
2216 http://www.themuseknoxville.org/ Knoxville\  
2217 Knoxville Botanical Garden and  
2218 Arboretum 35.982160 -83.881077 2743 Wimpole Ave, Knoxville, TN 37914, USA  
2219 (865) 862-8717

2214 <http://www.knoxgarden.org/> Knoxville\n Haunted Knoxville Ghost Tours  
2215 35.965963 -83.919553 36 Market Square #1404, Knoxville, TN 37902, USA (865)  
2216 377-9677  
2217 <http://www.hauntedknoxville.net/> Knoxville\n Three Rivers Rambler  
2218 35.952810 -83.940376 2560 University Commons Way, Knoxville, TN 37919, USA (865)  
2219 524-9411  
2220 <http://www.threeriversrambler.com/> Knoxville\n Charles Krutch Park  
2221 35.964414 -83.918695 504 Market St, Knoxville, TN 37902, USA (865)  
2222 215-4248  
2223 [https://www.knoxvilletn.gov/government/city\\_departments\\_offices/parks\\_and\\_recreation/parks/krutch\\_park](https://www.knoxvilletn.gov/government/city_departments_offices/parks_and_recreation/parks/krutch_park) Knoxville\n McClung Museum of Natural History & Culture  
2224 35.952005 -83.927209  
2225 1327 Cir Park Dr, Knoxville, TN 37996, USA (865) 974-2144  
2226 [http://mcclungmuseum](http://mcclungmuseum.utk.edu/)  
2227 .utk.edu/ Knoxville\n Knoxville Sightseeing 35.992790 -83.904128 2519 Mitchell St, Knoxville, TN  
2228 37917, USA (865) 566-0634  
2229 Unknown Knoxville\n Chilhowee Park & Exposition Center 35.996296 -83.884032  
2230 3301 E Magnolia Ave, Knoxville, TN 37914, USA (865) 215-1450  
2231 <https://chilhoweepark.com/> Knoxville\n Augusta Quarry  
2232 35.944772 -83.911886 3000 Fort Dickerson Rd SW, Knoxville, TN 37920, USA  
2233 Unknown  
2234 Unknown Knoxville\n Zoo Knoxville 35.999812 -83.888250  
2235 3500 Knoxville Zoo Dr, Knoxville, TN 37914, USA (865) 637-5331  
2236 <https://www.zooknoxville.org/> Knoxville\n Outdoor Knoxville Adventure Center  
2237 35.961902 -83.912663 900 Volunteer Landing Ln, Knoxville, TN 37915, USA (865)  
2238 228-8424  
2239 <http://www.outdoorknoxville.com> Knoxville\n James White Fort Association  
2240 35.962962 -83.912169 205 E Hill Ave, Knoxville, TN 37915, USA (865)  
2241 525-6514  
2242 <http://jameswhitesfort.org/> Knoxville\n Rowing Man Statue in Knoxville, TN  
2243 35.963611 -83.917519 W. Church Avenue &, S Gay St, Knoxville, TN 37902, USA  
2244 Unknown  
2245 Unknown Knoxville\n UT Gardens Knoxville 35.944000 -83.938260  
2246 2518 Jacob Dr, Knoxville, TN 37996, USA (865) 974-7324  
2247 <http://utgardens.tennessee.edu/locations/knoxville> Knoxville\n East Tennessee  
2248 Historical Society and Museum 35.964117 -83.917794 601 S Gay St, Knoxville, TN  
2249 37902, USA (865)  
2250 215-8830 [http://www.](http://www.easttnhistory.org/)  
2251 [easttnhistory.org/](http://www.easttnhistory.org/) Knoxville"}, {'Description': 'Restaurants in Knoxville', '  
2252 Content': '  
2253 Name Average Cost Cuisines Aggregate Rating  
2254 City\n Cafe Arabelle 29 French, BBQ, Desserts,  
2255 Seafood  
2256 3.6 Knoxville\n Les 3 Brasseurs 24 Desserts, Pizza, Italian,  
2257 Cafe, Indian 4.6 Knoxville\n Sky On 57 150  
2258 Tea,  
2259 Cafe, Pizza, Chinese, Seafood 3.4 Knoxville\n Cafe El Chico  
2260 67 Tea, French, Indian, Seafood 3.6  
2261 Knoxville\n  
2262 TcozY 85 Tea, Pizza, Mexican, Fast Food, Cafe 0.0  
2263 Knoxville\n Rama Vaishnav Bhojnalaya 46 Tea, Cafe, Pizza  
2264 , Desserts  
2265 3.0 Knoxville\n Burger Planet 45 Tea, Bakery,  
2266 American, Cafe 3.2 Knoxville\n Cafe Coffee Day 82  
2267 Fast  
2268 Food, American, BBQ, Italian 2.8 Knoxville\n Mamagoto  
2269 14 Indian, Mediterranean, Desserts, Seafood 4.1  
2270 Knoxville\n Ali Baba &  
2271 41 Dishes 70 Pizza, Desserts, Fast Food 3.5  
2272 Knoxville\n Punjabi Flavour 41 Tea, Bakery, Cafe,  
2273 Indian, Seafood  
2274 0.0 Knoxville\n Tandoori Tadka 23 Chinese, Pizza, Desserts,  
2275 Fast Food 0.0 Knoxville\n Coalition Cafe 35  
2276 Desserts,

2268 Tea, BBQ, Cafe, Mediterranean 3.4 Knoxville\n Chaat Corner  
2269 67 Cafe, Pizza 3.2  
2270 Knoxville\n  
2271 El Posto 63 Tea, Cafe, Pizza, BBQ 3.5  
2272 Knoxville\n La-Nawaab 17 French, Bakery,  
2273 Cafe, Seafood  
2274 3.0 Knoxville\n Peppers & Pipes 86 Chinese, Pizza,  
2275 Cafe, Italian 3.1 Knoxville\n Chit Chat 15  
2276 Pizza, Fast Food 3.1 Knoxville\n Anand Sweets 45  
2277 Fast Food, Pizza, Indian, Seafood 0.0 Knoxville\n  
2278 Open Kitchen  
2279 13 Pizza, Bakery, BBQ, Cafe, Indian, Mediterranean 3.0 Knoxville\n  
2280 Biryani By Kilo 17 Chinese, Pizza, Fast Food  
2281 4.1  
2282 Knoxville\n Chinese Food Corner 30 Indian, Desserts,  
2283 Fast Food 0.0 Knoxville\nThe Indian Kaffe Express 71  
2284 Tea, Desserts 3.8 Knoxville'}, {'Description': 'Accommodations in  
2285 Knoxville', 'Content': ' NAME price room  
2286 type  
2287 house\_rules minimum nights maximum occupancy review rate number city\n Cozy  
2288 Private Room in Chinatown/ Lower East Side 132.0 Private room  
2289 No pets &  
2290 No visitors 7.0 2 4.0 Knoxville\n  
2291 The Diamond Room 1008.0 Private room  
2292 No parties &  
2293 No visitors 1.0 1 5.0 Knoxville\n  
2294 Light-filled Room in Renovated Apt 310.0 Private room  
2295 No pets 2.0 2 2.0 Knoxville\n  
2296 Private Room 922.0 Private room  
2297 No  
2298 visitors 1.0 1 4.0 Knoxville\n  
2299 Beautiful & Private Manhattan Room 721.0 Private room No  
2300 parties & No  
2301 smoking 1.0 1 2.0 Knoxville\n Lg  
2302 Quiet Artist Home -Ditmas Park - 561.0 Entire home/apt No visitors & No  
2303 pets & No  
2304 smoking 24.0 4 2.0 Knoxville\n Beautiful 1-  
2305 bdrm apt in tranquil Inwood building 1091.0 Entire home/apt No  
2306 smoking &  
2307 No pets 2.0 4 4.0 Knoxville\n 1,600sq ft  
2308 modern duplex in new harlem brownstone 1159.0 Entire home/apt No smoking & No  
2309 children under 10 &  
2310 No pets 2.0 5 2.0 Knoxville\n  
2311 Brooklyn Sunny room 5 min to subway 793.0 Private room No visitors &  
2312 No children  
2313 under 10 2.0 2 5.0 Knoxville\n Amazing Large  
2314 Sunny Studio in Greenwich Village 144.0 Entire home/apt No  
2315 parties & No  
2316 smoking 30.0 3 4.0 Knoxville\n Private 1  
2317 Bdrm Suite in Historic Brownstone 479.0 Private room  
2318 No  
2319 visitors 2.0 2 2.0 Knoxville\n Charming  
2320 bedroom with huge terrace in Greenpoint 712.0 Private room No parties & No  
2321 children under 10 & No  
2322 visitors 2.0 1 3.0 Knoxville\n  
2323 Huge quiet modern oasis 1165.0 Entire home/apt  
2324 No smoking & No  
2325 parties 2.0 8 2.0 Knoxville\nPrivate large  
2326 room near LGA airport with queen bed 552.0 Private room  
2327 No pets 1.0 1 2.0 Knoxville\n 12 east 86  
2328 th Steet Upper East Side Building 632.0 Entire home/apt  
2329 No  
2330 smoking 30.0 3 2.0 Knoxville\n Spacious  
2331 Garden Apt in Clinton Hill/Ft Greene 1006.0 Entire home/apt No children under  
10 & No visitors & No

2322 smoking 3.0 3 3.0 Knoxville\n

2323 LES Penthouse 993.0 Entire home/apt No

2324 children under 10 & No

2325 parties 2.0 8 4.0 Knoxville'}, {'Description

2326 ': 'Attractions in Chattanooga', 'Content': "

2327 Name Latitude Longitude Address

2328 Phone

2328 Website City\n The Chattanooga Zoo at Warner Park

2329 35.042632 -85.282099 301 N Holtzclaw Ave, Chattanooga, TN 37404, USA (423)

2330 697-1322

2331 http://www.chattzoo.org/ Chattanooga\n Rock

2332 City Gardens 34.973392 -85.350177 1400 Patten Rd, Lookout Mountain, GA

2333 30750, USA (706)

2333 820-2531 http://seerockcity.

2334 com/?utm\_source=gmb&utm\_medium=organic Chattanooga\n

2335 Tennessee Aquarium 35.055823 -85.311065 1 Broad St, Chattanooga, TN

2336 37402, USA (423) 265-0695

2337 http://www.tnaqua.org/ Chattanooga\n Lookout Mountain

2338 Incline Railway 35.009546 -85.328564 3917 St Elmo Ave, Chattanooga, TN

2339 37409, USA (423)

2340 821-4224

2341 http://www.ridetheincline.com/ Chattanooga\n

2342 Coolidge Park 35.060506 -85.306857 150 River St, Chattanooga, TN 37405,

2343 USA (423) 643-6311

2344 http://www.chattanooga.gov/public-works/parks/directory-of-park-facilities

2345 Chattanooga\n Creative Discovery Museum 35.052700

2346 -85.312280

2347 321 Chestnut St, Chattanooga, TN 37402, USA (423) 756-2738

2348 http://www.cdmfun.org/

2349 Chattanooga\n Chattanooga Ghost Tours Inc

2350 35.051231 -85.309057 57 E 5th St, Chattanooga, TN 37402, USA (423)

2351 800-5998

2352 https://chattanoogaghosttours.com/ Chattanooga\n

2353 Ross's Landing 35.056741 -85.310302

2354 201 Riverfront Pkwy, Chattanooga, TN 37402,

2355 USA (423) 643-6311 http://www.chattanooga.gov

2356 /public-works/parks/directory-of-park-facilities Chattanooga\n

2357 Hunter Museum of American Art 35.055894 -85.306405 10 Bluff View Ave,

2358 Chattanooga, TN 37403, USA (423) 267-0968

2359 http://www.huntermuseum.org/ Chattanooga\n

2360 Walnut Street Bridge 35.058330 -85.307296 1 Walnut St, Chattanooga,

2361 TN 37403, USA

2362 (423) 643-6096 http://www.chattanooga.gov/

2363 public-works/parks/directory-of-park-facilities Chattanooga\n

2364 Tennessee Riverpark 35.095938 -85.245311 4301 Amnicola Hwy, Chattanooga, TN

2365 37402, USA (423) 209-5370

2366 http://parks.hamiltontn.gov/ Chattanooga\n Moccasin Bend National

2367 Archeological District 35.053024 -85.328656 301 Manufacturers Rd, Chattanooga,

2368 TN 37405, USA

2369 (423) 648-5623 https://www.nps.gov/chch/learn/historyculture/

2370 moccasin-bend-national-archeological-district.htm Chattanooga\n

2371 Ruby Falls 35.019084 -85.339355 1720 Scenic Hwy, Chattanooga, TN 37409, USA

2372 (423) 821-2544

2373 http://www.rubyfalls.com/ Chattanooga\n

2374 Cravens House 35.013782 -85.341495

2375 Point Park Visitor Ctr, Lookout Mountain, TN 37350, USA

(423) 821-7786

https://www.nps.gov/places/cravenshouselom.htm Chattanooga\n

2372 Williams Echo Dome 35.054215 -85.304836 430 E 2nd St, Chattanooga, TN

2373 37403, USA Unknown

2374 Unknown Chattanooga\n

2375 Reflection Riding Arboretum & Nature

Center 35.010130 -85.364585 400 Garden Rd, Chattanooga, TN 37419, USA

(423) 821-1160

2376 https://www.reflectionriding.org/ Chattanooga\n

2377 Audubon Acres 34.999350 -85.180340

2378 900 N Sanctuary Rd, Chattanooga, TN 37421,

2379 USA (423) 892-1499

2380 http

2381 ://www.chattanoogaaudubon.org/ Chattanooga\n

2382 Sculpture Fields at Montague Park 35.026175 -85.293624 1800 Polk St,

2383 Chattanooga, TN 37408, USA (423) 266-7288

2384 http://www.sculpturefields.org/ Chattanooga\nBessie Smith Cultural Center and

2385 Chattanooga African American Museum 35.044634 -85.306467 200 E M L King Blvd,

2386 Chattanooga, TN 37403,

2387 USA (423) 266-8658 http://www.bessiesmithcc.org/?utm\_source=google&utm\_medium=

2388 wix\_google\_business\_profile&utm\_campaign=3948799342172393035 Chattanooga\n

2389 Wooden Horse 35.055836 -85.306713 10 Bluff View Ave, Chattanooga, TN 37403,

2390 USA Unknown

2391 Unknown Chattanooga"}, {'Description': 'Restaurants in Chattanooga', 'Content': "

2392 Name Average Cost Cuisines

2393 Aggregate

2394 Rating City\n P.F. Chang's 33

2395 Cafe, Bakery, BBQ, Seafood 4.2 Chattanooga\n L'amandier

2396 28 Cafe, BBQ, Mediterranean, Seafood 4.3 Chattanooga\n

2397 Anupam Sweet 82 Desserts,

2398 American, BBQ

2399 3.4 Chattanooga\n Nirula's 64

2400 French, Desserts, Seafood 2.6 Chattanooga\n Anupam Sweets &

2401 Restaurant 62

2402 Tea, Cafe, Desserts, Seafood 3.0 Chattanooga\n Bisque Bakery

2403 90 French, Desserts, Seafood 3.9

2404 Chattanooga\n

2405 Nooba 93 Tea, Mexican, Bakery, Seafood 3.8

2406 Chattanooga\n The Royal 39 French,

2407 Pizza,

2408 Bakery, Fast Food 3.3 Chattanooga\n Curry n Phulka 40

2409 Tea, Chinese, Fast Food 2.8 Chattanooga\n

2410 DCK-

2411 Dana Choga's Kitchen 67 Cafe, Pizza, Mediterranean

2412 3.5 Chattanooga\n Chaayos 54

2413 Desserts, Cafe, BBQ, Chinese, Seafood 3.3 Chattanooga\n

2414 Truffles 53 Cafe, Bakery, BBQ, Fast Food, Chinese, American

2415 3.2

2416 Chattanooga\n Liquid 15 Tea, Pizza, Bakery,

2417 Mediterranean, Seafood 4.0 Chattanooga\n Tpot

2418 14

2419 Desserts, Tea, Pizza, Mexican, Cafe, Indian 0.0 Chattanooga\n Sardar A

2420 Pure Meat Shop 24 Bakery, Pizza, American

2421 3.4 Chattanooga\n Warehouse Cafe 90 Fast

2422 Food, Pizza, Seafood 3.7 Chattanooga\n The Beer Cafe

2423 62

2424 Tea, Pizza, Fast Food 3.8 Chattanooga\n Subway

2425 66 Tea, Cafe, Bakery, Desserts 2.4 Chattanooga

2426 \n

2427 Indi-QUE 57 Tea, Italian, BBQ, Fast Food, Cafe

2428 3.3 Chattanooga\nEssex Collections Patisserie 48

2429 Chinese, BBQ, Fast Food 3.4 Chattanooga\n Food Adda

100 Desserts, Pizza, Fast Food, Mediterranean, Seafood 0.0

2422 Chattanooga\n

2423 Bikanervala 65 Cafe, Bakery

2424 3.2 Chattanooga\n Havmor Ice Cream 75

2425 Pizza, Bakery, Fast

2426 Food, Chinese, Seafood 3.6 Chattanooga\n Moti Mahal Delux

2427 52 Tea, Pizza, BBQ, Cafe, Mediterranean 2.7

2428 Chattanooga\n

2429 Aggarwal Sweet India 30 Tea, Desserts

2.8 Chattanooga\n Jack Po!tato's 63

Tea,

2430 Bakery, Mediterranean, Seafood 3.3 Chattanooga Muradabadi  
2431 16 Cafe, Pizza, Bakery, BBQ 3.0  
2432 Chattanooga Vishal Restaurant 27 Tea,  
2433 Pizza, BBQ, Seafood 3.1 Chattanooga Raj Sweets  
2434 69  
2435 Desserts, Tea, BBQ, Cafe, Mediterranean 2.8 Chattanooga Punjabee's  
2436 Darbar 41 Tea, French, Bakery, Fast Food, Chinese  
2437 3.0  
2438 Chattanooga Basil Tree 20 Bakery, Indian,  
2439 American, Desserts 2.4 Chattanooga Domino's Pizza  
2440 77  
2441 Bakery, Pizza, American, BBQ 2.8 Chattanooga Pizza Hut Delivery  
2442 25 Tea, Cafe, American, Indian 3.9  
2443 Chattanooga  
2444 Ichiban 87 Mexican, Bakery, Desserts  
2445 3.9 Chattanooga Habibi 95  
2446 Cafe, BBQ, Seafood 3.7 Chattanooga Cook Du Kdu 69  
2447 Pizza, Fast Food, Cafe, Indian, Mediterranean, Seafood 3.1 Chattanooga  
2448 Go Krazy 64 French, BBQ, Fast Food  
2449 3.5 Chattanooga Bentoya 41 Tea,  
2450 Fast  
2451 Food, Bakery, Seafood 3.8 Chattanooga Burger's King  
2452 16 Tea, Bakery, Desserts 4.1 Chattanooga  
2453 "},  
2454 {'Description': 'Accommodations in Chattanooga', 'Content': '  
2455 NAME price room type  
2456 house\_rules minimum  
2457 nights maximum occupancy review rate number city Affordable Private  
2458 Spacious Room in Brooklyn 790.0 Private room No  
2459 parties  
2460 2.0 2 3.0 Chattanooga  
2461 Sunny One Bedroom 722.0 Entire home/apt No pets & No children under 10  
2462 1.0  
2463 5 3.0 Chattanooga Upper West / Morningside Heights Apt, Near Subway  
2464 290.0 Entire home/apt No visitors 3.0  
2465 3  
2466 3.0 Chattanooga Sunny room+Pvte office in huge loft 728.0 Private room  
2467 No parties 4.0 2  
2468 5.0  
2469 Chattanooga Extra Cozy Room in Center of Williamsburg 1033.0 Private room  
2470 No pets 1.0 1  
2471 1.0  
2472 Chattanooga Luxury & Charm. Steps from Christopher Park! 301.0 Entire home/apt  
2473 No pets & No visitors & No parties 5.0 2  
2474 4.0  
2475 Chattanooga Fort Greene Room 722.0 Private room  
2476 No visitors & No children under 10 2.0 1  
2477 2.0  
2478 Chattanooga 973.0  
2479 Entire home/apt No pets 10.0  
2480 6  
2481 4.0 Chattanooga 4 bdrm/2 bath apt. Central Pk, Columbia U. 564.0 Entire home/  
2482 apt No visitors & No pets 30.0 4  
2483 3.0  
2484 Chattanooga Cozy room in Bushwick- 15 min to the city 933.0 Private room  
2485 No parties & No pets 1.0 2  
2486 2.0  
2487 Chattanooga Modern apartment w/ gorgeous view 921.0 Private room  
2488 No pets 3.0 2  
2489 1.0  
2490 Chattanooga Artsy Private BR in Fort Greene Cumberland 727.0 Private room  
2491 No visitors 30.0 1  
2492 3.0  
2493

2484 Chattanooga\n Studio Deluxe 1 - Wyndham Midtown 45 848.0 Private room No  
2485 pets & No smoking & No children under 10 3.0 1  
2486 2.0  
2487 Chattanooga\n Trendy and Stylish Downtown apartment !!! 110.0 Entire home/apt  
2488 No pets 3.0 2  
2489 2.0  
2490 Chattanooga\n Sonder | 180 Water | Luxurious 1BR + Rooftop 676.0 Entire home/apt  
2491 No pets & No smoking 29.0 5  
2492 3.0  
2492 Chattanooga'}, {'Description': 'Flight from Houston to Nashville on 2022-03-21', '  
2493 Content': 'Flight Number Price DepTime ArrTime ActualElapsedTime FlightDate  
2494 OriginCityName DestCityName  
2495 Distance\n F3827247 241 14:22 16:10 1 hours 48 minutes 2022-03-21 Houston  
2496 Nashville 657.0\n F3827724 147 18:35 20:33 1 hours 58 minutes  
2497 2022-03-21  
2498 Houston Nashville 657.0\n F3827820 157 12:07 14:02 1 hours 55 minutes  
2499 2022-03-21 Houston Nashville 657.0\n F3827930 305 07:36 09:41 2  
2500 hours 5  
2501 minutes 2022-03-21 Houston Nashville 657.0\n F3956407 264 11:38  
2502 13:15 1 hours 37 minutes 2022-03-21 Houston Nashville 670.0\n  
2503 F3956408 246  
2504 08:08 09:47 1 hours 39 minutes 2022-03-21 Houston Nashville 670.0\n  
2505 F3956409 145 17:36 19:12 1 hours 36 minutes 2022-03-21 Houston  
2506 Nashville 670.0\n  
2507 F3956532 225 07:14 08:52 1 hours 38 minutes 2022-03-21 Houston Nashville  
2508 657.0\n F3956533 292 11:42 13:19 1 hours 37 minutes 2022-03-21  
2509 Houston  
2510 Nashville 657.0\n F3956534 248 18:20 20:06 1 hours 46 minutes 2022-03-21  
2511 Houston Nashville 657.0\n F4038756 272 10:11 11:52 1 hours 41  
2512 minutes  
2513 2022-03-21 Houston Nashville 657.0\n F4038966 149 16:33 18:13 1  
2514 hours 40 minutes 2022-03-21 Houston Nashville 657.0'}, {'Description': '  
2515 Self-driving  
2516 from Houston to Nashville', 'Content': 'self-driving, from Houston to Nashville,  
2517 duration: 11 hours 50 mins, distance: 1,253 km, cost: 62'}, {'Description': '  
2518 Taxi from Houston to  
2519 Nashville', 'Content': 'taxi, from Houston to Nashville, duration: 11 hours 50 mins,  
2520 distance: 1,253 km, cost: 1253'}, {'Description': 'Flight from Nashville to  
2521 Knoxville on 2022-03-23',  
2522 'Content': 'There is no flight from Nashville to Knoxville on 2022-03-23.'}, {'  
2523 Description': 'Self-driving from Nashville to Knoxville', 'Content': 'self-  
2524 driving, from Nashville to  
2525 Knoxville, duration: 2 hours 42 mins, distance: 290 km, cost: 14'}, {'Description':  
2526 'Taxi from Nashville to Knoxville', 'Content': 'taxi, from Nashville to  
2527 Knoxville, duration: 2 hours 42  
2528 mins, distance: 290 km, cost: 290'}, {'Description': 'Flight from Knoxville to  
2529 Chattanooga on 2022-03-25', 'Content': 'There is no flight from Knoxville to  
2530 Chattanooga on 2022-03-25.'},  
2531 {'Description': 'Self-driving from Knoxville to Chattanooga', 'Content': 'self-  
2532 driving, from Knoxville to Chattanooga, duration: 1 hour 41 mins, distance: 180  
2533 km, cost: 9'}, {'Description':  
2534 'Taxi from Knoxville to Chattanooga', 'Content': 'taxi, from Knoxville to  
2535 Chattanooga, duration: 1 hour 41 mins, distance: 180 km, cost: 180'}, {'  
2536 Description': 'Flight from Chattanooga to  
2537 Houston on 2022-03-27', 'Content': 'There is no flight from Chattanooga to Houston  
2538 on 2022-03-27.'}, {'Description': 'Self-driving from Chattanooga to Houston', '  
2539 Content': 'self-driving,  
2540 from Chattanooga to Houston, duration: 11 hours 47 mins, distance: 1,309 km, cost:  
2541 65'}, {'Description': 'Taxi from Chattanooga to Houston', 'Content': 'taxi,  
2542 from Chattanooga to Houston,  
2543 duration: 11 hours 47 mins, distance: 1,309 km, cost: 1309'}]  
2544 Query: Could you design a one-week travel itinerary for two, departing from Houston  
2545 and touring three cities in Tennessee from March 21st to March 27th, 2022? Our  
2546 budget is now \$8,200. We

```

2538 require accommodations that allow smoking and should ideally be private rooms. As
2539 for transportation, we would prefer not to self-drive.
2540 Travel Plan (please only output the JSON string without explanatory information):
2541
2542 =====RESPONSE OF o1-preview=====
2543
2544 [
2545   {
2546     "day": 1,
2547     "people_number": 2,
2548     "current_city": "from Houston to Nashville",
2549     "transportation": "Flight Number: F3827820, from Houston to Nashville,
2550       Departure Time: 12:07, Arrival Time: 14:02",
2551     "breakfast": "-",
2552     "attraction": "Country Music Hall of Fame and Museum, Nashville",
2553     "lunch": "Cafe Arabelle, Nashville",
2554     "dinner": "Les 3 Brasseurs, Nashville",
2555     "accommodation": "FiDi Cozy room overlooking East River, Nashville"
2556   },
2557   {
2558     "day": 2,
2559     "people_number": 2,
2560     "current_city": "Nashville",
2561     "transportation": "-",
2562     "breakfast": "Burger Planet, Nashville",
2563     "attraction": "Johnny Cash Museum; Centennial Park, Nashville",
2564     "lunch": "Mamagoto, Nashville",
2565     "dinner": "Oh! Calcutta, Nashville",
2566     "accommodation": "FiDi Cozy room overlooking East River, Nashville"
2567   },
2568   {
2569     "day": 3,
2570     "people_number": 2,
2571     "current_city": "from Nashville to Knoxville",
2572     "transportation": "Taxi, from Nashville to Knoxville, Departure Time: 09:00,
2573       Arrival Time: 11:42, Duration: 2 hours 42 minutes, Cost: $290",
2574     "breakfast": "Cafe El Chico, Nashville",
2575     "attraction": "World's Fair Park, Knoxville",
2576     "lunch": "Ali Baba & 41 Dishes, Knoxville",
2577     "dinner": "Sky On 57, Knoxville",
2578     "accommodation": "Light-filled Room in Renovated Apt, Knoxville"
2579   },
2580   {
2581     "day": 4,
2582     "people_number": 2,
2583     "current_city": "Knoxville",
2584     "transportation": "-",
2585     "breakfast": "Chaat Corner, Knoxville",
2586     "attraction": "Ijams Nature Center; Sunsphere, Knoxville",
2587     "lunch": "Anand Sweets, Knoxville",
2588     "dinner": "La-Nawaab, Knoxville",
2589     "accommodation": "Light-filled Room in Renovated Apt, Knoxville"
2590   },
2591   {
2592     "day": 5,
2593     "people_number": 2,
2594     "current_city": "from Knoxville to Chattanooga",
2595     "transportation": "Taxi, from Knoxville to Chattanooga, Departure Time:
2596       09:00, Arrival Time: 10:41, Duration: 1 hour 41 minutes, Cost: $180",
2597     "breakfast": "Anand Sweets, Knoxville",
2598     "attraction": "Tennessee Aquarium, Chattanooga",
2599     "lunch": "Anupam Sweet, Chattanooga",
2600     "dinner": "Bisque Bakery, Chattanooga",
2601     "accommodation": "Trendy and Stylish Downtown apartment !!!, Chattanooga"
2602   }
2603 ]

```



```

2592     {
2593         "day": 6,
2594         "people_number": 2,
2595         "current_city": "Chattanooga",
2596         "transportation": "-",
2597         "breakfast": "Anupam Sweets & Restaurant, Chattanooga",
2598         "attraction": "Rock City Gardens; Ruby Falls, Chattanooga",
2599         "lunch": "Chaayos, Chattanooga",
2600         "dinner": "L'amandier, Chattanooga",
2601         "accommodation": "Trendy and Stylish Downtown apartment !!!, Chattanooga"
2602     },
2603     {
2604         "day": 7,
2605         "people_number": 2,
2606         "current_city": "from Chattanooga to Houston",
2607         "transportation": "Taxi, from Chattanooga to Houston, Departure Time: 08:00,
2608             Arrival Time: 19:47, Duration: 11 hours 47 minutes, Cost: $1309",
2609         "breakfast": "Bisque Bakery, Chattanooga",
2610         "attraction": "Coolidge Park, Chattanooga",
2611         "lunch": "Nooba, Chattanooga",
2612         "dinner": "-",
2613         "accommodation": "-"
2614     }
2615 ]
2616 =====FEEDBACK FROM METACONTROLLER=====
2617 By inspecting your plan we find the following issue. Please refine your plan
2618 according to the feedback below:
2619 1. The room type should be private room :-> Trendy and Stylish Downtown apartment
2620 !!!
2621 2. The restaurant in day 5 breakfast is repeated.
2622 3. The accommodation Trendy and Stylish Downtown apartment !!!, Chattanooga do not
2623 obey the minimum nights rule.
2624 4. The lunch in day 1 is invalid or not in the data provided.
2625
2626 O.2 BACK PROMPT
2627
2628 You are a proficient planner. Based on the provided information and query, please
2629 give me a detailed plan, including specifics such as flight numbers (e.g.,
2630 F0123456), restaurant names, and
2631 accommodation names. Note that all the information in your plan should be derived
2632 from the provided data. You should give a travel plan in JSON format as shown
2633 in the example below.
2634 Additionally, all details should align with commonsense. The symbol '-' indicates
2635 that information is unnecessary. For example, in the provided sample, you do
2636 not need to plan after
2637 returning to the departure city. When you travel to two cities in one day, you
2638 should note it in the 'current_city' section as in the example (i.e., from A to
2639 B).
2640
2641 ***** Example *****
2642 Query: Could you create a travel plan for 7 people from Ithaca to Charlotte spanning
2643 3 days, from March 8th to March 14th, 2022, with a budget of $30,200?
2644 Travel Plan:
2645 [
2646     {
2647         "day": 1,
2648         "people_number": 7,
2649         "current_city": "from Ithaca to Charlotte",
2650         "transportation": "Flight Number: F3633413, from Ithaca to Charlotte,
2651             Departure Time: 05:38, Arrival Time: 07:46",
2652         "breakfast": "Nagaland's Kitchen, Charlotte",
2653         "attraction": "The Charlotte Museum of History, Charlotte",
2654         "lunch": "Cafe Maple Street, Charlotte",
2655         "dinner": "Bombay Vada Pav, Charlotte",

```

```

2646     "accommodation": "Affordable Spacious Refurbished Room in Bushwick!,
2647         Charlotte"
2648 },
2649 {
2650     "day": 2,
2651     "people_number": 7,
2652     "current_city": "Charlotte",
2653     "transportation": "-",
2654     "breakfast": "Olive Tree Cafe, Charlotte",
2655     "attraction": "The Mint Museum, Charlotte;Romare Bearden Park, Charlotte",
2656     "lunch": "Birbal Ji Dhaba, Charlotte",
2657     "dinner": "Pind Balluchi, Charlotte",
2658     "accommodation": "Affordable Spacious Refurbished Room in Bushwick!,
2659         Charlotte"
2660 },
2661 {
2662     "day": 3,
2663     "people_number": 7,
2664     "current_city": "from Charlotte to Ithaca",
2665     "transportation": "Flight Number: F3786167, from Charlotte to Ithaca,
2666         Departure Time: 21:42, Arrival Time: 23:26",
2667     "breakfast": "Books Monument, Charlotte",
2668     "attraction": "Books Monument, Charlotte",
2669     "lunch": "Olive Tree Cafe, Charlotte",
2670     "dinner": "Kylin Skybar, Charlotte",
2671     "accommodation": "-"
2672 }
2673 ]
2674 ***** Example Ends *****
2675
2676 Given information: [{"Description": 'Attractions in Nashville', 'Content': '
2677     Name Latitude Longitude
2678     Address
2679     Phone
2680     Website    City\nCountry Music Hall of Fame and Museum 36.158263 -86.776126 222 Rep.
2681     John Lewis Way S, Nashville, TN 37203, USA (615) 416-2001
2682     https://countrymusichalloffame.org/ Nashville\n    Nashville Zoo at Grassmere
2683     36.089705 -86.742096 3777 Nolensville Pk, Nashville, TN 37211, USA (615)
2684     833-1534
2685     http://www.nashvillezoo.org/ Nashville\n Belle Meade Historic Site & Winery
2686     36.104916 -86.864695 5025 Harding Pike, Nashville, TN 37205, USA (615) 356-0501
2687     https://visitbellemeade.com/ Nashville\n    Johnny Cash Museum 36.160939
2688     -86.775757    119 3rd Ave S, Nashville, TN 37201, USA (615) 256-1777
2689     http://www.johnnycashmuseum.com/ Nashville\n    Centennial Park
2690     36.148946 -86.812750    2500 West End Ave, Nashville, TN 37203, USA (615)
2691     862-8400
2692     https://www.nashville.gov/Parks-and-Recreation/Parks/Centennial-Park.aspx Nashville\n
2693     n    Grand Ole Opry 36.206857 -86.692108    600 Opry Mills Dr,
2694     Nashville, TN
2695     37214, USA (615) 871-6779
2696     https://www.opry.com/?utm_campaign=opry&utm_medium=organicsearch&utm_source=
2697     googlemybusiness&utm_audience=tofu_googlemybusiness&utm_content=
2698     brandstory_google_my_business_website_link
2699     Nashville\n    Frist Art Museum 36.157897 -86.783853    919
2700     Broadway, Nashville, TN 37203, USA (615) 244-3340
2701     https://fristartmuseum.org/ Nashville\n    The Parthenon 36.149674
2702     -86.813347    2500 West End Ave, Nashville, TN 37203, USA (615) 862-8431
2703     https://www.nashvilleparthenon.com/ Nashville\n    Nashville Shores
2704     36.158209 -86.604915    4001 Bell Rd, Hermitage, TN 37076, USA (615)
2705     889-7050
2706     http://www.nashvilleshores.com/lodging Nashville\n Musicians Hall of Fame and Museum
2707     36.167668 -86.782399    401 Gay St, Nashville, TN 37219, USA (615) 244-3263
2708     https://www.musicianshalloffame.com/ Nashville\n Bicentennial Capitol Mall State
2709     Park 36.170887 -86.787589 600 James Robertson Pkwy, Nashville, TN 37243, USA
2710     (888) 867-2757

```

2700 <https://tnstateparks.com/parks/bicentennial-mall> Nashville\n Lane Motor  
2701 Museum 36.140197 -86.734580 702 Murfreesboro Pike, Nashville, TN 37210, USA  
2702 (615) 742-7445  
2703 <http://www.lanemotormuseum.org/> Nashville\n Honky Tonk Highway 36.160393  
2704 -86.778432 501 Broadway, Nashville, TN 37203, USA (800) 657-6910  
2705 <https://www.visitmusiccity.com/visitors/honkytonkhighway> Nashville\n Adventure  
2706 Science Center 36.146614 -86.775481 800 Fort Negley Blvd, Nashville, TN 37203,  
2707 USA (615)  
2708 862-5160  
2708 <http://www.adventuresci.org/> Nashville\n Patsy Cline Museum 36.160904  
2709 -86.776030 119 3rd Ave S, Nashville, TN 37201, USA (615) 454-4722  
2710 <https://www.patsymuseum.com/> Nashville\n Andrew Jacksons Hermitage 36.213757  
2711 -86.615310 4580 Rachels Ln, Hermitage, TN 37076, USA (615) 889-2941  
2712 <http://www.thehermitage.com/> Nashville\n NashTrash Tours 36.169474  
2713 -86.788254 900 Rosa L Parks Blvd, Nashville, TN 37208, USA (615) 226-7300  
2714 <http://www.nashtrash.com/> Nashville\n Fort Nashborough 36.164268  
2715 -86.775430 170 1st Ave N, Nashville, TN 37201, USA (615) 862-8400  
2716 <https://www.nashville.gov/Parks-and-Recreation/Historic-Sites/Fort-Nashborough.aspx>  
2717 Nashville\n Ryman Auditorium 36.161248 -86.778471 116 5th Ave N,  
2718 Nashville, TN 37219, USA (615) 889-3060  
2719 [https://www.ryman.com/?utm\\_campaign=ryman&utm\\_medium=organicsearch&utm\\_source=](https://www.ryman.com/?utm_campaign=ryman&utm_medium=organicsearch&utm_source=googlemybusiness&utm_audience=tofu_googlemybusiness&utm_content=brandstory_google_my_business_website_link)  
2720 [googlemybusiness&utm\\_audience=tofu\\_googlemybusiness&utm\\_content=](https://www.ryman.com/?utm_campaign=ryman&utm_medium=organicsearch&utm_source=googlemybusiness&utm_audience=tofu_googlemybusiness&utm_content=brandstory_google_my_business_website_link)  
2721 [brandstory\\_google\\_my\\_business\\_website\\_link](https://www.ryman.com/?utm_campaign=ryman&utm_medium=organicsearch&utm_source=brandstory_google_my_business_website_link)  
2722 Nashville\n Nashville Public Square Park 36.166611 -86.778126 Union St & 3rd  
2723 Ave N, Nashville, TN 37201, USA (615) 743-3090  
2724 <https://www.nashvilledowntown.com/go/public-square-park> Nashville'}, {'Description':  
2725 'Restaurants in Nashville', 'Content': "  
2726 Cuisines Aggregate Rating City\n Name Average Cost  
2727 Bangkok 1 44  
2728 Cafe, Seafood 3.3 Nashville\n  
2729 Bablu Fast  
2730 Food 13 Desserts, Pizza, Mexican, BBQ, Fast Food 0.0 Nashville\  
2731 n Full Dabba 77 Pizza, Mexican, Fast Food, Cafe,  
2732 American  
2733 0.0 Nashville\  
2734 Chinese, Seafood 4.5 Nashville\  
2735 35  
2736 Tea, Mexican, Seafood 0.0 Nashville\  
2737 The Toddy Shop  
2738 92 Desserts, Fast Food 3.4 Nashville\  
2739 GoGourmet 34 Tea, Cafe, Indian, Mediterranean 3.7  
2740 Nashville\  
2741 Govinda's Confectionery 15 Tea, Bakery,  
2742 Mediterranean, Fast Food  
2743 3.1 Nashville\  
2744 Kitchen King 13 Pizza, Italian, BBQ,  
2745 Cafe, Mediterranean 0.0 Nashville\  
2746 55  
2747 Tea, Seafood 3.8 Nashville\  
2748 Smoke House Deli 59  
2749 Cafe, Mediterranean, Seafood 4.0 Nashville\  
2750 Dialogue  
2751 Lounge & Caf 83 Tea, Indian, BBQ, Cafe, American, Seafood 3.0  
2752 Nashville\  
2753 Meenakshi Bhawan 24  
2754 Tea,  
2755 Desserts 3.1 Nashville\  
2756 Oh! Calcutta 83  
2757 Tea, Mexican, BBQ, Cafe, Indian 4.4 Nashville\  
2758 Chicago  
2759 Pizza 20 Tea, French, Pizza, BBQ 3.2 Nashville\  
2760 n Malhotra Restaurant 76  
2761 Seafood  
2762 2.9 Nashville\  
2763 Madras Cafe 88 Pizza,  
2764 Desserts, Seafood 0.0 Nashville\  
2765 nDakshin - Sheraton New Delhi Hotel  
2766 63  
2767 Cafe, Pizza, Indian, Bakery 4.0 Nashville\  
2768 1911 Bar - The  
2769 Imperial 76 Desserts, Pizza, Italian, American, Seafood 3.2  
2770 Nashville\  
2771 Punjabi Virsa 35 Desserts, Italian, BBQ, Cafe, Seafood 0.0  
2772 Nashville\  
2773 Sagar Dhaba 55 Tea, Bakery, BBQ,  
2774 Cafe,

2754 Mediterranean 0.0 Nashville\n Kettle & Kegs 60  
2755 Tea, Seafood 0.0 Nashville\n  
2756 Kargo 84 Tea, Pizza, Desserts 4.4 Nashville  
2757 "}, {'Description': 'Accommodations in Nashville', 'Content': '  
2758 NAME price room type house\_rules minimum nights  
2759 maximum occupancy review rate number city\nHuge 2 Bedroom, Great Location,  
2760 Express Metro  
2761 745.0 Entire home/apt No smoking & No children under 10 & No pets 4.0  
2762 4 5.0 Nashville\n Clean and large bedroom in a  
2763 private house 474.0  
2764 Private room No smoking 1.0 1  
2765 4.0 Nashville\n Brooklyn Heights gem 993.0  
2766 Entire  
2767 home/apt No pets 2.0 6  
2768 5.0 Nashville\n Lovely room in heart of Williamsburg 61.0  
2769 Private room  
2770 No pets 2.0 1 4.0 Nashville\n \*Light & Love\*  
2771 vibrant, historic, sleeps 4 679.0 Entire home/apt No  
2772 pets  
2773 5.0 2 2.0 Nashville\n FiDi Cozy room overlooking  
2774 East River 870.0 Private room No parties  
2775 1.0  
2776 2 5.0 Nashville\n Charming 1BR with sun-nook in Brooklyn 256.0  
2777 Entire home/apt No children under 10 & No pets & No smoking 3.0  
2778 2  
2779 4.0 Nashville\n Cozy bedroom close to Manhattan 576.0 Private room  
2780 No children under 10 1.0 1  
2781 3.0  
2782 Nashville'}, {'Description': 'Attractions in Knoxville', 'Content': "  
2783 Name Latitude Longitude  
2784 Address  
2785 Phone  
2786 Website City\n World's Fair Park 35.962577 -83.924192  
2787 525 Henley St, Knoxville, TN 37902, USA (865) 215-1158  
2788 http://  
2789 worldsfairpark.org/ Knoxville\n  
2790 Knoxville Museum of Art 35.962426 -83.925229 1050 Worlds Fair Park Dr, Knoxville, TN  
2791 37916, USA (865) 525-6101  
2792 http://www.knoxart.org/ Knoxville\n Sunsphere 35.961707  
2793 -83.923353 810 Clinch Ave, Knoxville, TN 37902, USA (865) 314-0660  
2794 http://www.sunspheretickets.com/ Knoxville\n Ijams Nature Center  
2795 35.956454 -83.866775 2915 Island Home Ave, Knoxville, TN 37920, USA (865)  
2796 577-4717  
2797 http://www.ijams.org/ Knoxville\n Knoxville Walking Tours 35.966448  
2798 -83.919167 301 S Gay St, Knoxville, TN 37902, USA (865) 309-4522  
2799 http://knoxvillewalkingtours.com/ Knoxville\n Muse Knoxville  
2800 35.997617 -83.885467 516 N Beaman St, Knoxville, TN 37914, USA (865)  
2801 594-1494  
2802 http://www.themuseknoxville.org/ Knoxville\n Knoxville Botanical Garden and  
2803 Arboretum 35.982160 -83.881077 2743 Wimpole Ave, Knoxville, TN 37914, USA  
2804 (865) 862-8717  
2805 http://www.knoxgarden.org/ Knoxville\n Haunted Knoxville Ghost Tours  
2806 35.965963 -83.919553 36 Market Square #1404, Knoxville, TN 37902, USA (865)  
2807 377-9677  
2808 http://www.hauntedknoxville.net/ Knoxville\n Three Rivers Rambler  
2809 35.952810 -83.940376 2560 University Commons Way, Knoxville, TN 37919, USA (865)  
2810 524-9411  
2811 http://www.threeriversrambler.com/ Knoxville\n Charles Krutch Park  
2812 35.964414 -83.918695 504 Market St, Knoxville, TN 37902, USA (865)  
2813 215-4248  
2814 https://www.knoxvilletn.gov/government/city\_departments\_offices/parks\_and\_recreation  
2815 /parks/krutch\_park Knoxville\n McClung Museum of Natural History & Culture  
2816 35.952005 -83.927209

2808 1327 Cir Park Dr, Knoxville, TN 37996, USA (865) 974-2144  
2809 <http://mcclungmuseum>  
2810 .utk.edu/ Knoxville\n  
2811 Knoxville Sightseeing 35.992790 -83.904128 2519 Mitchell St, Knoxville, TN  
2812 37917, USA (865) 566-0634  
2813 Unknown Knoxville\n Chilhowee Park & Exposition Center 35.996296 -83.884032  
2814 3301 E Magnolia Ave, Knoxville, TN 37914, USA (865) 215-1450  
2815 <https://chilhoweepark.com/> Knoxville\n Augusta Quarry  
2816 35.944772 -83.911886 3000 Fort Dickerson Rd SW, Knoxville, TN 37920, USA  
2817 Unknown  
2818 Unknown Knoxville\n Zoo Knoxville 35.999812 -83.888250  
2819 3500 Knoxville Zoo Dr, Knoxville, TN 37914, USA (865) 637-5331  
2820 <https://www.zooknoxville.org/> Knoxville\n Outdoor Knoxville Adventure Center  
2821 35.961902 -83.912663 900 Volunteer Landing Ln, Knoxville, TN 37915, USA (865)  
2822 228-8424  
2823 <http://www.outdoorknoxville.com> Knoxville\n James White Fort Association  
2824 35.962962 -83.912169 205 E Hill Ave, Knoxville, TN 37915, USA (865)  
2825 525-6514  
2826 <http://jameswhitesfort.org/> Knoxville\n Rowing Man Statue in Knoxville, TN  
2827 35.963611 -83.917519 W. Church Avenue &, S Gay St, Knoxville, TN 37902, USA  
2828 Unknown  
2829 Unknown Knoxville\n UT Gardens Knoxville 35.944000 -83.938260  
2830 2518 Jacob Dr, Knoxville, TN 37996, USA (865) 974-7324  
2831 <http://utgardens.tennessee.edu/locations/knoxville> Knoxville\n East Tennessee  
2832 Historical Society and Museum 35.964117 -83.917794 601 S Gay St, Knoxville, TN  
2833 37902, USA (865)  
2834 215-8830 <http://www.easttnhistory.org/> Knoxville"}, {'Description': 'Restaurants in Knoxville', 'Content': '  
2835 Name Average Cost Cuisines Aggregate Rating  
2836 City\n Cafe Arabelle 29 French, BBQ, Desserts,  
2837 Seafood  
2838 3.6 Knoxville\n Les 3 Brasseurs 24 Desserts, Pizza, Italian,  
2839 Cafe, Indian 4.6 Knoxville\n Sky On 57 150  
2840 Tea,  
2841 Cafe, Pizza, Chinese, Seafood 3.4 Knoxville\n Cafe El Chico  
2842 67 Tea, French, Indian, Seafood 3.6  
2843 Knoxville\n  
2844 TcozY 85 Tea, Pizza, Mexican, Fast Food, Cafe 0.0  
2845 Knoxville\n Rama Vaishnav Bhojnalaya 46 Tea, Cafe, Pizza  
2846 , Desserts  
2847 3.0 Knoxville\n Burger Planet 45 Tea, Bakery,  
2848 American, Cafe 3.2 Knoxville\n Cafe Coffee Day 82  
2849 Fast  
2850 Food, American, BBQ, Italian 2.8 Knoxville\n Mamagoto  
2851 14 Indian, Mediterranean, Desserts, Seafood 4.1  
2852 Knoxville\n Ali Baba &  
2853 41 Dishes 70 Pizza, Desserts, Fast Food 3.5  
2854 Knoxville\n Punjabi Flavour 41 Tea, Bakery, Cafe,  
2855 Indian, Seafood  
2856 0.0 Knoxville\n Tandoori Tadka 23 Chinese, Pizza, Desserts,  
2857 Fast Food 0.0 Knoxville\n Coalition Cafe 35  
2858 Desserts,  
2859 Tea, BBQ, Cafe, Mediterranean 3.4 Knoxville\n Chaat Corner  
2860 67 Cafe, Pizza 3.2  
2861 Knoxville\n  
2862 El Posto 63 Tea, Cafe, Pizza, BBQ 3.5  
2863 Knoxville\n La-Nawaab 17 French, Bakery,  
2864 Cafe, Seafood  
2865 3.0 Knoxville\n Peppers & Pipes 86 Chinese, Pizza,  
2866 Cafe, Italian 3.1 Knoxville\n Chit Chat 15  
2867 Pizza, Fast Food 3.1 Knoxville\n Anand Sweets 45  
2868 Fast Food, Pizza, Indian, Seafood 0.0 Knoxville\n  
2869 Open Kitchen

2862 13 Pizza, Bakery, BBQ, Cafe, Indian, Mediterranean 3.0 Knoxville\n

2863 Biryani By Kilo 17 Chinese, Pizza, Fast Food

2864 4.1

2865 Knoxville\n Chinese Food Corner 30 Indian, Desserts,

2866 Fast Food 0.0 Knoxville\nThe Indian Kaffe Express 71

2867 Tea, Desserts 3.8 Knoxville'}, {'Description': 'Accommodations in

2868 Knoxville', 'Content': ' NAME price room

2869 type

2869 house\_rules minimum nights maximum occupancy review rate number city\n Cozy

2870 Private Room in Chinatown/ Lower East Side 132.0 Private room

2871 No pets &

2872 No visitors 7.0 2 4.0 Knoxville\n

2873 The Diamond Room 1008.0 Private room

2874 No parties &

2874 No visitors 1.0 1 5.0 Knoxville\n

2875 Light-filled Room in Renovated Apt 310.0 Private room

2876 No pets 2.0 2 2.0 Knoxville\n

2877 Private Room 922.0 Private room

2878 No

2878 visitors 1.0 1 4.0 Knoxville\n

2879 Beautiful & Private Manhattan Room 721.0 Private room No

2880 parties & No

2881 smoking 1.0 1 2.0 Knoxville\n Lg

2882 Quiet Artist Home -Ditmas Park - 561.0 Entire home/apt No visitors & No

2883 pets & No

2883 smoking 24.0 4 2.0 Knoxville\n Beautiful 1-

2884 bdrm apt in tranquil Inwood building 1091.0 Entire home/apt No

2885 smoking &

2886 No pets 2.0 4 4.0 Knoxville\n 1,600sq ft

2887 modern duplex in new harlem brownstone 1159.0 Entire home/apt No smoking & No

2888 children under 10 &

2888 No pets 2.0 5 2.0 Knoxville\n

2889 Brooklyn Sunny room 5 min to subway 793.0 Private room No visitors &

2890 No children

2891 under 10 2.0 2 5.0 Knoxville\n Amazing Large

2892 Sunny Studio in Greenwich Village 144.0 Entire home/apt No

2893 parties & No

2893 smoking 30.0 3 4.0 Knoxville\n Private 1

2894 Bdrm Suite in Historic Brownstone 479.0 Private room

2895 No

2896 visitors 2.0 2 2.0 Knoxville\n Charming

2897 bedroom with huge terrace in Greenpoint 712.0 Private room No parties & No

2898 children under 10 & No

2899 visitors 2.0 1 3.0 Knoxville\n

2899 Huge quiet modern oasis 1165.0 Entire home/apt

2900 No smoking & No

2901 parties 2.0 8 2.0 Knoxville\nPrivate large

2902 room near LGA airport with queen bed 552.0 Private room

2903 No pets 1.0 1 2.0 Knoxville\n 12 east 86

2904 th Steet Upper East Side Building 632.0 Entire home/apt

2905 No

2905 smoking 30.0 3 2.0 Knoxville\n Spacious

2906 Garden Apt in Clinton Hill/Ft Greene 1006.0 Entire home/apt No children under

2907 10 & No visitors & No

2908 smoking 3.0 3 3.0 Knoxville\n

2908 LES Penthouse 993.0 Entire home/apt No

2909 children under 10 & No

2910 parties 2.0 8 4.0 Knoxville'}, {'Description

2911 ': 'Attractions in Chattanooga', 'Content': "

2912 Name Latitude Longitude Address

2913 Phone

2913 Website City\n The Chattanooga Zoo at Warner Park

2914 35.042632 -85.282099 301 N Holtzclaw Ave, Chattanooga, TN 37404, USA (423)

2915 697-1322

2916 <http://www.chattzoo.org/> Chattanooga\n Rock  
 2917 City Gardens 34.973392 -85.350177 1400 Patten Rd, Lookout Mountain, GA  
 2918 30750, USA (706)  
 2919 820-2531 [http://seerockcity.com/?utm\\_source=gmb&utm\\_medium=organic](http://seerockcity.com/?utm_source=gmb&utm_medium=organic) Chattanooga\n  
 2920 Tennessee Aquarium 35.055823 -85.311065 1 Broad St, Chattanooga, TN  
 2921 37402, USA (423) 265-0695  
 2922 <http://www.tnaqua.org/> Chattanooga\n Lookout Mountain  
 2923 Incline Railway 35.009546 -85.328564 3917 St Elmo Ave, Chattanooga, TN  
 2924 37409, USA (423)  
 2925 821-4224  
 2926 <http://www.ridetheincline.com/> Chattanooga\n  
 2927 Coolidge Park 35.060506 -85.306857 150 River St, Chattanooga, TN 37405,  
 2928 USA (423) 643-6311  
 2929 <http://www.chattanooga.gov/public-works/parks/directory-of-park-facilities>  
 2930 Chattanooga\n Creative Discovery Museum 35.052700  
 2931 -85.312280  
 2932 321 Chestnut St, Chattanooga, TN 37402, USA (423) 756-2738  
 2933 <http://www.cdmfun.org/>  
 2934 Chattanooga\n Chattanooga Ghost Tours Inc  
 2935 35.051231 -85.309057 57 E 5th St, Chattanooga, TN 37402, USA (423)  
 2936 800-5998  
 2937 <https://chattanoogaghosttours.com/> Chattanooga\n  
 2938 Ross's Landing 35.056741 -85.310302  
 2939 201 Riverfront Pkwy, Chattanooga, TN 37402,  
 2940 USA (423) 643-6311 <http://www.chattanooga.gov/public-works/parks/directory-of-park-facilities> Chattanooga\n  
 2941 Hunter Museum of American Art 35.055894 -85.306405 10 Bluff View Ave,  
 2942 Chattanooga, TN 37403, USA (423) 267-0968  
 2943 <http://www.huntermuseum.org/> Chattanooga\n  
 2944 Walnut Street Bridge 35.058330 -85.307296 1 Walnut St, Chattanooga,  
 2945 TN 37403, USA  
 2946 (423) 643-6096 <http://www.chattanooga.gov/public-works/parks/directory-of-park-facilities> Chattanooga\n  
 2947 Tennessee Riverpark 35.095938 -85.245311 4301 Amnicola Hwy, Chattanooga, TN  
 2948 37402, USA (423) 209-5370  
 2949 <http://parks.hamiltontn.gov/> Chattanooga\n Moccasin Bend National  
 2950 Archeological District 35.053024 -85.328656 301 Manufacturers Rd, Chattanooga,  
 2951 TN 37405, USA  
 2952 (423) 648-5623 <https://www.nps.gov/chch/learn/historyculture/moccasin-bend-national-archeological-district.htm> Chattanooga\n  
 2953 Ruby Falls 35.019084 -85.339355 1720 Scenic Hwy, Chattanooga, TN 37409, USA  
 2954 (423) 821-2544  
 2955 <http://www.rubyfalls.com/> Chattanooga\n  
 2956 Cravens House 35.013782 -85.341495  
 2957 Point Park Visitor Ctr, Lookout Mountain, TN 37350, USA  
 2958 (423) 821-7786  
 2959 <https://www.nps.gov/places/cravenshouselom.htm> Chattanooga\n  
 2960 Williams Echo Dome 35.054215 -85.304836 430 E 2nd St, Chattanooga, TN  
 2961 37403, USA Unknown  
 2962 Unknown Chattanooga\n Reflection Riding Arboretum & Nature  
 2963 Center 35.010130 -85.364585 400 Garden Rd, Chattanooga, TN 37419, USA  
 2964 (423) 821-1160  
 2965 <https://www.reflectionriding.org/> Chattanooga\n  
 2966 Audubon Acres 34.999350 -85.180340  
 2967 900 N Sanctuary Rd, Chattanooga, TN 37421,  
 2968 USA (423) 892-1499  
 2969 <http://www.chattanoogaaudubon.org/> Chattanooga\n  
 2970 Sculpture Fields at Montague Park 35.026175 -85.293624 1800 Polk St,  
 2971 Chattanooga, TN 37408, USA (423) 266-7288  
 2972

2970 <http://www.sculpturefields.org/> Chattanooga\nBessie Smith Cultural Center and  
2971 Chattanooga African American Museum 35.044634 -85.306467 200 E M L King Blvd,  
2972 Chattanooga, TN 37403,  
2973 USA (423) 266-8658 [http://www.bessiesmithcc.org/?utm\\_source=google&utm\\_medium=](http://www.bessiesmithcc.org/?utm_source=google&utm_medium=wix_google_business_profile&utm_campaign=3948799342172393035)  
2974 [wix\\_google\\_business\\_profile&utm\\_campaign=3948799342172393035](http://www.bessiesmithcc.org/?utm_source=google&utm_medium=wix_google_business_profile&utm_campaign=3948799342172393035) Chattanooga\  
2975 Wooden Horse 35.055836 -85.306713 10 Bluff View Ave, Chattanooga, TN 37403,  
2976 USA Unknown  
2977 Unknown Chattanooga"}, {'Description': 'Restaurants in Chattanooga', 'Content': "  
2978 Aggregate Name Average Cost Cuisines  
2979 Rating City\n P.F. Chang's 33  
2980 Cafe, Bakery, BBQ, Seafood 4.2 Chattanooga\n L'amandier  
2981 28 Cafe, BBQ, Mediterranean, Seafood 4.3 Chattanooga\  
2982 Anupam Sweet 82 Desserts,  
2983 American, BBQ  
2984 3.4 Chattanooga\n Nirula's 64  
2985 French, Desserts, Seafood 2.6 Chattanooga\n Anupam Sweets &  
2986 Restaurant 62  
2987 Tea, Cafe, Desserts, Seafood 3.0 Chattanooga\n Bisque Bakery  
2988 90 French, Desserts, Seafood 3.9  
2989 Chattanooga\  
2990 Nooba 93 Tea, Mexican, Bakery, Seafood 3.8  
2991 Chattanooga\n The Royal 39 French,  
2992 Pizza,  
2993 Bakery, Fast Food 3.3 Chattanooga\n Curry n Phulka 40  
2994 Tea, Chinese, Fast Food 2.8 Chattanooga\  
2995 DCK-  
2996 Dana Choga's Kitchen 67 Cafe, Pizza, Mediterranean  
2997 3.5 Chattanooga\  
2998 Chaayos 54  
2999 Desserts, Cafe, BBQ, Chinese, Seafood 3.3 Chattanooga\  
3000 Truffles 53 Cafe, Bakery, BBQ, Fast Food, Chinese, American  
3001 3.2  
3002 Chattanooga\n Liquid 15 Tea, Pizza, Bakery,  
3003 Mediterranean, Seafood 4.0 Chattanooga\n Tpot  
3004 14  
3005 Desserts, Tea, Pizza, Mexican, Cafe, Indian 0.0 Chattanooga\  
3006 Pure Meat Shop 24 Bakery, Pizza, American  
3007 3.4 Chattanooga\n Warehouse Cafe 90 Fast  
3008 Food, Pizza, Seafood 3.7 Chattanooga\  
3009 The Beer Cafe  
3010 62  
3011 Tea, Pizza, Fast Food 3.8 Chattanooga\  
3012 66 Subway  
3013 \n Tea, Cafe, Bakery, Desserts 2.4 Chattanooga  
3014 Indi-QUE 57 Tea, Italian, BBQ, Fast Food, Cafe  
3015 3.3 Chattanooga\nEssex Collections Patisserie 48  
3016 Chinese, BBQ, Fast Food 3.4 Chattanooga\  
3017 Food Adda  
3018 100 Desserts, Pizza, Fast Food, Mediterranean, Seafood 0.0  
3019 Chattanooga\  
3020 Bikanervala 65 Cafe, Bakery  
3021 3.2 Chattanooga\  
3022 Havmor Ice Cream 75  
3023 Pizza, Bakery, Fast  
3024 Food, Chinese, Seafood 3.6 Chattanooga\  
3025 Moti Mahal Delux  
3026 52 Tea, Pizza, BBQ, Cafe, Mediterranean 2.7  
3027 Chattanooga\  
3028 Aggarwal Sweet India 30 Tea, Desserts  
3029 2.8 Chattanooga\  
3030 Jack Po!tato's 63  
3031 Tea,  
3032 Bakery, Mediterranean, Seafood 3.3 Chattanooga\  
3033 Muradabadi  
3034 16 Cafe, Pizza, Bakery, BBQ 3.0  
3035 Chattanooga\  
3036 Vishal Restaurant 27 Tea,  
3037 Pizza, BBQ, Seafood 3.1 Chattanooga\  
3038 Raj Sweets  
3039 69  
3040 Desserts, Tea, BBQ, Cafe, Mediterranean 2.8 Chattanooga\  
3041 Punjabee's  
3042 Darbar 41 Tea, French, Bakery, Fast Food, Chinese  
3043 3.0



3024 Chattanooga\n Basil Tree 20 Bakery, Indian,  
3025 American, Desserts 2.4 Chattanooga\n Domino's Pizza  
3026 77  
3027 Bakery, Pizza, American, BBQ 2.8 Chattanooga\n Pizza Hut Delivery  
3028 25 Tea, Cafe, American, Indian 3.9  
3029 Chattanooga\n  
3029 Ichiban 87 Mexican, Bakery, Desserts  
3030 3.9 Chattanooga\n Habibi 95  
3031 Cafe, BBQ, Seafood 3.7 Chattanooga\n Cook Du Kdu 69  
3032 Pizza, Fast Food, Cafe, Indian, Mediterranean, Seafood 3.1 Chattanooga\n  
3033 Go Krazy 64 French, BBQ, Fast Food  
3034 3.5 Chattanooga\n Bentoya 41 Tea,  
3035 Fast  
3035 Food, Bakery, Seafood 3.8 Chattanooga\n Burger's King  
3036 16 Tea, Bakery, Desserts 4.1 Chattanooga  
3037 "],  
3038 {'Description': 'Accommodations in Chattanooga', 'Content': '  
3039 NAME price room type  
3040 house\_rules minimum  
3040 nights maximum occupancy review rate number city\n Affordable Private  
3041 Spacious Room in Brooklyn 790.0 Private room No  
3042 parties  
3043 2.0 2 3.0 Chattanooga\n  
3044 Sunny One Bedroom 722.0 Entire home/apt No pets & No children under 10  
3045 1.0  
3045 5 3.0 Chattanooga\nUpper West / Morningside Heights Apt, Near Subway  
3046 290.0 Entire home/apt No visitors 3.0  
3047 3  
3048 3.0 Chattanooga\n Sunny room+Pvte office in huge loft 728.0 Private room  
3049 No parties 4.0 2  
3050 5.0  
3050 Chattanooga\n Extra Cozy Room in Center of Williamsburg 1033.0 Private room  
3051 No pets 1.0 1  
3052 1.0  
3053 Chattanooga\n Luxury & Charm. Steps from Christopher Park! 301.0 Entire home/apt  
3054 No pets & No visitors & No parties 5.0 2  
3055 4.0  
3055 Chattanooga\n Fort Greene Room 722.0 Private room  
3056 No visitors & No children under 10 2.0 1  
3057 2.0  
3058 Chattanooga\n 973.0  
3059 Entire home/apt No pets 10.0  
3060 6  
3060 4.0 Chattanooga\n 4 bdrm/2 bath apt. Central Pk, Columbia U. 564.0 Entire home/  
3061 apt No visitors & No pets 30.0 4  
3062 3.0  
3063 Chattanooga\n Cozy room in Bushwick- 15 min to the city 933.0 Private room  
3064 No parties & No pets 1.0 2  
3065 2.0  
3065 Chattanooga\n Modern apartment w/ gorgeous view 921.0 Private room  
3066 No pets 3.0 2  
3067 1.0  
3068 Chattanooga\n Artsy Private BR in Fort Greene Cumberland 727.0 Private room  
3069 No visitors 30.0 1  
3070 3.0  
3070 Chattanooga\n Studio Deluxe 1 - Wyndham Midtown 45 848.0 Private room No  
3071 pets & No smoking & No children under 10 3.0 1  
3072 2.0  
3073 Chattanooga\n Trendy and Stylish Downtown apartment !!! 110.0 Entire home/apt  
3074 No pets 3.0 2  
3075 2.0  
3075 Chattanooga\n Sonder | 180 Water | Luxurious 1BR + Rooftop 676.0 Entire home/apt  
3076 No pets & No smoking 29.0 5  
3077 3.0

3078 Chattanooga'}, {'Description': 'Flight from Houston to Nashville on 2022-03-21', '  
3079 Content': 'Flight Number Price DepTime ArrTime ActualElapsedTime FlightDate  
3080 OriginCityName DestCityName  
3081 Distance\n F3827247 241 14:22 16:10 1 hours 48 minutes 2022-03-21 Houston  
3082 Nashville 657.0\n F3827724 147 18:35 20:33 1 hours 58 minutes  
3083 2022-03-21  
3084 Houston Nashville 657.0\n F3827820 157 12:07 14:02 1 hours 55 minutes  
3085 2022-03-21 Houston Nashville 657.0\n F3827930 305 07:36 09:41 2  
3086 hours 5  
3087 minutes 2022-03-21 Houston Nashville 657.0\n F3956407 264 11:38  
3088 13:15 1 hours 37 minutes 2022-03-21 Houston Nashville 670.0\n  
3089 F3956408 246  
3090 08:08 09:47 1 hours 39 minutes 2022-03-21 Houston Nashville 670.0\n  
3091 F3956409 145 17:36 19:12 1 hours 36 minutes 2022-03-21 Houston  
3092 Nashville 670.0\n  
3093 F3956532 225 07:14 08:52 1 hours 38 minutes 2022-03-21 Houston Nashville  
3094 657.0\n F3956533 292 11:42 13:19 1 hours 37 minutes 2022-03-21  
3095 Houston  
3096 Nashville 657.0\n F3956534 248 18:20 20:06 1 hours 46 minutes 2022-03-21  
3097 Houston Nashville 657.0\n F4038756 272 10:11 11:52 1 hours 41  
3098 minutes  
3099 2022-03-21 Houston Nashville 657.0\n F4038966 149 16:33 18:13 1  
3100 hours 40 minutes 2022-03-21 Houston Nashville 657.0'}, {'Description': '  
3101 Self-driving  
3102 from Houston to Nashville', 'Content': 'self-driving, from Houston to Nashville,  
3103 duration: 11 hours 50 mins, distance: 1,253 km, cost: 62'}, {'Description': '  
3104 Taxi from Houston to  
3105 Nashville', 'Content': 'taxi, from Houston to Nashville, duration: 11 hours 50 mins,  
3106 distance: 1,253 km, cost: 1253'}, {'Description': 'Flight from Nashville to  
3107 Knoxville on 2022-03-23',  
3108 'Content': 'There is no flight from Nashville to Knoxville on 2022-03-23.'}, {'  
3109 Description': 'Self-driving from Nashville to Knoxville', 'Content': 'self-  
3110 driving, from Nashville to  
3111 Knoxville, duration: 2 hours 42 mins, distance: 290 km, cost: 14'}, {'Description':  
3112 'Taxi from Nashville to Knoxville', 'Content': 'taxi, from Nashville to  
3113 Knoxville, duration: 2 hours 42  
3114 mins, distance: 290 km, cost: 290'}, {'Description': 'Flight from Knoxville to  
3115 Chattanooga on 2022-03-25', 'Content': 'There is no flight from Knoxville to  
3116 Chattanooga on 2022-03-25.'},  
3117 {'Description': 'Self-driving from Knoxville to Chattanooga', 'Content': 'self-  
3118 driving, from Knoxville to Chattanooga, duration: 1 hour 41 mins, distance: 180  
3119 km, cost: 9'}, {'Description':  
3120 'Taxi from Knoxville to Chattanooga', 'Content': 'taxi, from Knoxville to  
3121 Chattanooga, duration: 1 hour 41 mins, distance: 180 km, cost: 180'}, {'  
3122 Description': 'Flight from Chattanooga to  
3123 Houston on 2022-03-27', 'Content': 'There is no flight from Chattanooga to Houston  
3124 on 2022-03-27.'}, {'Description': 'Self-driving from Chattanooga to Houston', '  
3125 Content': 'self-driving,  
3126 from Chattanooga to Houston, duration: 11 hours 47 mins, distance: 1,309 km, cost:  
3127 65'}, {'Description': 'Taxi from Chattanooga to Houston', 'Content': 'taxi,  
3128 from Chattanooga to Houston,  
3129 duration: 11 hours 47 mins, distance: 1,309 km, cost: 1309'}]  
3130 Query: Could you design a one-week travel itinerary for two, departing from Houston  
3131 and touring three cities in Tennessee from March 21st to March 27th, 2022? Our  
budget is now \$8,200. We  
require accommodations that allow smoking and should ideally be private rooms. As  
for transportation, we would prefer not to self-drive.  
Travel Plan:  
[  
{  
"day": 1,  
"people\_number": 2,  
"current\_city": "from Houston to Nashville",  
"transportation": "Flight Number: F3827820, from Houston to Nashville,  
Departure Time: 12:07, Arrival Time: 14:02",

```

3132     "breakfast": "-",
3133     "attraction": "Country Music Hall of Fame and Museum, Nashville",
3134     "lunch": "Cafe Arabelle, Nashville",
3135     "dinner": "Les 3 Brasseurs, Nashville",
3136     "accommodation": "FiDi Cozy room overlooking East River, Nashville"
3137 },
3138 {
3139     "day": 2,
3140     "people_number": 2,
3141     "current_city": "Nashville",
3142     "transportation": "-",
3143     "breakfast": "Burger Planet, Nashville",
3144     "attraction": "Johnny Cash Museum; Centennial Park, Nashville",
3145     "lunch": "Mamagoto, Nashville",
3146     "dinner": "Oh! Calcutta, Nashville",
3147     "accommodation": "FiDi Cozy room overlooking East River, Nashville"
3148 },
3149 {
3150     "day": 3,
3151     "people_number": 2,
3152     "current_city": "from Nashville to Knoxville",
3153     "transportation": "Taxi, from Nashville to Knoxville, Departure Time: 09:00,
3154         Arrival Time: 11:42, Duration: 2 hours 42 minutes, Cost: $290",
3155     "breakfast": "Cafe El Chico, Nashville",
3156     "attraction": "World's Fair Park, Knoxville",
3157     "lunch": "Ali Baba & 41 Dishes, Knoxville",
3158     "dinner": "Sky On 57, Knoxville",
3159     "accommodation": "Light-filled Room in Renovated Apt, Knoxville"
3160 },
3161 {
3162     "day": 4,
3163     "people_number": 2,
3164     "current_city": "Knoxville",
3165     "transportation": "-",
3166     "breakfast": "Chaat Corner, Knoxville",
3167     "attraction": "Ijams Nature Center; Sunsphere, Knoxville",
3168     "lunch": "Anand Sweets, Knoxville",
3169     "dinner": "La-Nawaab, Knoxville",
3170     "accommodation": "Light-filled Room in Renovated Apt, Knoxville"
3171 },
3172 {
3173     "day": 5,
3174     "people_number": 2,
3175     "current_city": "from Knoxville to Chattanooga",
3176     "transportation": "Taxi, from Knoxville to Chattanooga, Departure Time:
3177         09:00, Arrival Time: 10:41, Duration: 1 hour 41 minutes, Cost: $180",
3178     "breakfast": "Anand Sweets, Knoxville",
3179     "attraction": "Tennessee Aquarium, Chattanooga",
3180     "lunch": "Anupam Sweet, Chattanooga",
3181     "dinner": "Bisque Bakery, Chattanooga",
3182     "accommodation": "Trendy and Stylish Downtown apartment !!!, Chattanooga"
3183 },
3184 {
3185     "day": 6,
3186     "people_number": 2,
3187     "current_city": "Chattanooga",
3188     "transportation": "-",
3189     "breakfast": "Anupam Sweets & Restaurant, Chattanooga",
3190     "attraction": "Rock City Gardens; Ruby Falls, Chattanooga",
3191     "lunch": "Chaayos, Chattanooga",
3192     "dinner": "L'amandier, Chattanooga",
3193     "accommodation": "Trendy and Stylish Downtown apartment !!!, Chattanooga"
3194 },
3195 {
3196     "day": 7,

```

```

3186     "people_number": 2,
3187     "current_city": "from Chattanooga to Houston",
3188     "transportation": "Taxi, from Chattanooga to Houston, Departure Time: 08:00,
3189         Arrival Time: 19:47, Duration: 11 hours 47 minutes, Cost: $1309",
3190     "breakfast": "Bisque Bakery, Chattanooga",
3191     "attraction": "Coolidge Park, Chattanooga",
3192     "lunch": "Nooba, Chattanooga",
3193     "dinner": "-",
3194     "accommodation": "-"
3195 ]

```

```

3196 By inspecting your plan we find the following issue. Please refine your plan
3197 according to the feedback below:
3198 1. The room type should be private room :-> Trendy and Stylish Downtown apartment
3199    !!!
3200 2. The restaurant in day 5 breakfast is repeated.
3201 3. The accommodation Trendy and Stylish Downtown apartment !!!, Chattanooga do not
3202    obey the minumum nights rule.
3203 4. The lunch in day 1 is invalid or not in the data provided.

```

```

3204 Travel Plan (please only output the JSON string without explanatory information):

```

```

3205 =====RESPONSE OF o1-preview=====
3206

```

```

3207 [
3208   {
3209     "day": 1,
3210     "people_number": 2,
3211     "current_city": "from Houston to Nashville",
3212     "transportation": "Flight Number: F3827820, from Houston to Nashville,
3213         Departure Time: 12:07, Arrival Time: 14:02",
3214     "breakfast": "-",
3215     "attraction": "Country Music Hall of Fame and Museum, Nashville",
3216     "lunch": "Smoke House Deli, Nashville",
3217     "dinner": "Les 3 Brasseurs, Nashville",
3218     "accommodation": "FiDi Cozy room overlooking East River, Nashville"
3219   },
3220   {
3221     "day": 2,
3222     "people_number": 2,
3223     "current_city": "Nashville",
3224     "transportation": "-",
3225     "breakfast": "Burger Planet, Nashville",
3226     "attraction": "Johnny Cash Museum; Centennial Park, Nashville",
3227     "lunch": "Mamagoto, Nashville",
3228     "dinner": "Oh! Calcutta, Nashville",
3229     "accommodation": "FiDi Cozy room overlooking East River, Nashville"
3230   },
3231   {
3232     "day": 3,
3233     "people_number": 2,
3234     "current_city": "from Nashville to Knoxville",
3235     "transportation": "Taxi, from Nashville to Knoxville, Departure Time: 09:00,
3236         Arrival Time: 11:42, Duration: 2 hours 42 minutes, Cost: $290",
3237     "breakfast": "Cafe El Chico, Nashville",
3238     "attraction": "World's Fair Park, Knoxville",
3239     "lunch": "Ali Baba & 41 Dishes, Knoxville",
3240     "dinner": "Sky On 57, Knoxville",
3241     "accommodation": "Light-filled Room in Renovated Apt, Knoxville"
3242   },
3243   {
3244     "day": 4,
3245     "people_number": 2,
3246     "current_city": "Knoxville",

```

```

3240     "transportation": "-",
3241     "breakfast": "Chaat Corner, Knoxville",
3242     "attraction": "Ijams Nature Center; Sunsphere, Knoxville",
3243     "lunch": "Anand Sweets, Knoxville",
3244     "dinner": "La-Nawaab, Knoxville",
3245     "accommodation": "Light-filled Room in Renovated Apt, Knoxville"
3246   },
3247   {
3248     "day": 5,
3249     "people_number": 2,
3250     "current_city": "from Knoxville to Chattanooga",
3251     "transportation": "Taxi, from Knoxville to Chattanooga, Departure Time:
3252       09:00, Arrival Time: 10:41, Duration: 1 hour 41 minutes, Cost: $180",
3253     "breakfast": "Burger Planet, Knoxville",
3254     "attraction": "Tennessee Aquarium, Chattanooga",
3255     "lunch": "Anupam Sweet, Chattanooga",
3256     "dinner": "Bisque Bakery, Chattanooga",
3257     "accommodation": "Affordable Private Spacious Room in Brooklyn, Chattanooga"
3258   },
3259   {
3260     "day": 6,
3261     "people_number": 2,
3262     "current_city": "Chattanooga",
3263     "transportation": "-",
3264     "breakfast": "Anupam Sweets & Restaurant, Chattanooga",
3265     "attraction": "Rock City Gardens; Ruby Falls, Chattanooga",
3266     "lunch": "Chaayos, Chattanooga",
3267     "dinner": "L'amandier, Chattanooga",
3268     "accommodation": "Affordable Private Spacious Room in Brooklyn, Chattanooga"
3269   },
3270   {
3271     "day": 7,
3272     "people_number": 2,
3273     "current_city": "from Chattanooga to Houston",
3274     "transportation": "Taxi, from Chattanooga to Houston, Departure Time: 08:00,
3275       Arrival Time: 19:47, Duration: 11 hours 47 minutes, Cost: $1309",
3276     "breakfast": "Bisque Bakery, Chattanooga",
3277     "attraction": "Coolidge Park, Chattanooga",
3278     "lunch": "Nooba, Chattanooga",
3279     "dinner": "-",
3280     "accommodation": "-"
3281   }
3282 ]

```

## 3278 P PROMPTS FOR TRIP PLANNING

### 3280 P.1 FIRST ITERATION

```

3282 You are an expert at planning trips. You are given a few constraints regarding the
3283 cities to visit and the durations of staying at each city. You are also given
3284 the flight information between the cities('and' suggests flights are possible
3285 both ways and 'from - to -' suggests
3286 one way flights). You can travel to any city only once.
3287 Produce only one plan, strictly adhering to the format shown in the examples below,
3288 without any extra information. Your answer should start with 'SOLUTION:'.
3289 Here are a few example tasks and solutions:
3290 TASK: You plan to visit 10 European cities for 21 days in total. You only take
3291 direct flights to commute between cities. You plan to stay in Vienna for 3 days.
3292 You want to spend 5 days in Frankfurt. You want to spend 2 days in Oslo. You
3293 are going to attend a wedding in Oslo
3294 between day 20 and day 21. You want to spend 3 days in Prague. You would like to
3295 visit Valencia for 2 days. You want to meet a friend in Valencia between day 17

```

3294 and day 18. You plan to stay in Dubrovnik for 2 days. You would like to visit  
3295 Edinburgh for 5 days. From day 1 to  
3296 day 5, there is a annual show you want to attend in Edinburgh. You plan to stay in  
3297 London for 2 days. You plan to visit relatives in London between day 12 and day  
3298 13. You would like to visit Munich for 3 days. You would like to meet your  
3299 friends at Munich between day 18 and  
3300 day 20 to tour together. You would like to visit Budapest for 3 days.

3301 Here are the cities that have direct flights:  
3302 Valencia and Munich, Vienna and Munich, Vienna and Valencia, London and Budapest,  
3303 London and Oslo, Edinburgh and Budapest, Frankfurt and Budapest, Frankfurt and  
3304 London, Prague and Oslo, Edinburgh and Oslo, Edinburgh and Munich, Prague and  
3305 Munich, London and Prague, Edinburgh  
3306 and London, Edinburgh and Frankfurt, Dubrovnik and Munich, Dubrovnik and Vienna,  
3307 Munich and Oslo, Dubrovnik and Oslo, Budapest and Munich, Frankfurt and Prague,  
3308 Vienna and London, Frankfurt and Vienna, Frankfurt and Oslo, Frankfurt and  
3309 Munich, Vienna and Oslo, Vienna and  
3310 Prague, Budapest and Oslo, Budapest and Prague, London and Valencia, London and  
3311 Munich, Frankfurt and Dubrovnik, Prague and Valencia, Frankfurt and Valencia,  
3312 Edinburgh and Prague, Vienna and Budapest.

3312 Find a trip plan of visiting the cities for 21 days by taking direct flights to  
3313 commute between them.

3314 SOLUTION: Here is the trip plan for visiting the 10 European cities for 21 days:

3315 \*\*Day 1-5:\*\* Arriving in Edinburgh and visit Edinburgh for 5 days.

3316 \*\*Day 5:\*\* Fly from Edinburgh to Frankfurt.

3317 \*\*Day 5-9:\*\* Visit Frankfurt for 5 days.

3318 \*\*Day 9:\*\* Fly from Frankfurt to Dubrovnik.

3319 \*\*Day 9-10:\*\* Visit Dubrovnik for 2 days.

3320 \*\*Day 10:\*\* Fly from Dubrovnik to Vienna.

3321 \*\*Day 10-12:\*\* Visit Vienna for 3 days.

3322 \*\*Day 12:\*\* Fly from Vienna to London.

3323 \*\*Day 12-13:\*\* Visit London for 2 days.

3324 \*\*Day 13:\*\* Fly from London to Budapest.

3325 \*\*Day 13-15:\*\* Visit Budapest for 3 days.

3326 \*\*Day 15:\*\* Fly from Budapest to Prague.

3327 \*\*Day 15-17:\*\* Visit Prague for 3 days.

3328 \*\*Day 17:\*\* Fly from Prague to Valencia.

3329 \*\*Day 17-18:\*\* Visit Valencia for 2 days.

3330 \*\*Day 18:\*\* Fly from Valencia to Munich.

3331 \*\*Day 18-20:\*\* Visit Munich for 3 days.

3332 \*\*Day 20:\*\* Fly from Munich to Oslo.

3333 \*\*Day 20-21:\*\* Visit Oslo for 2 days.

3332 TASK: You plan to visit 10 European cities for 22 days in total. You only take  
3333 direct flights to commute between cities. You want to spend 4 days in Helsinki.

3334 You plan to stay in Vilnius for 2 days. You plan to stay in Istanbul for 5  
3335 days. From day 17 to day 21, there is a

3336 annual show you want to attend in Istanbul. You want to spend 2 days in Hamburg. You  
3337 plan to stay in Porto for 2 days. During day 11 and day 12, you have to attend  
3338 a conference in Porto. You want to spend 4 days in Athens. You would like to  
3339 visit Prague for 2 days. You plan

3340 to visit relatives in Prague between day 13 and day 14. You want to spend 2 days in  
3341 Frankfurt. You want to spend 3 days in Krakow. You want to spend 5 days in  
3342 Munich. You would like to meet your friends at Munich between day 1 and day 5  
3343 to tour together.

3343 Here are the cities that have direct flights:  
3344 from Krakow to Vilnius, Helsinki and Hamburg, Hamburg and Athens, Munich and  
3345 Frankfurt, Hamburg and Porto, Munich and Istanbul, Prague and Athens, Frankfurt  
3346 and Athens, Munich and Athens, Munich and Prague, from Vilnius to Munich,  
3347 Hamburg and Istanbul, Frankfurt and  
3348 Istanbul, Munich and Krakow, Munich and Hamburg, Munich and Helsinki, Prague and  
3349 Istanbul, Frankfurt and Vilnius, Helsinki and Istanbul, Athens and Vilnius,

3348 Krakow and Frankfurt, Helsinki and Frankfurt, Porto and Frankfurt, Frankfurt  
3349 and Prague, Istanbul and Vilnius, Krakow  
3350 and Istanbul, Krakow and Prague, Munich and Porto, Helsinki and Vilnius, Helsinki  
3351 and Prague, Porto and Istanbul, Hamburg and Frankfurt, Krakow and Helsinki,  
3352 Athens and Istanbul.

3353 Find a trip plan of visiting the cities for 22 days by taking direct flights to  
3354 commute between them.

3355 SOLUTION: Here is the trip plan for visiting the 10 European cities for 22 days:  
3356

3357 \*\*Day 1-5:\*\* Arriving in Munich and visit Munich for 5 days.  
3358 \*\*Day 5:\*\* Fly from Munich to Krakow.  
3359 \*\*Day 5-7:\*\* Visit Krakow for 3 days.  
3360 \*\*Day 7:\*\* Fly from Krakow to Helsinki.  
3361 \*\*Day 7-10:\*\* Visit Helsinki for 4 days.  
3362 \*\*Day 10:\*\* Fly from Helsinki to Hamburg.  
3363 \*\*Day 10-11:\*\* Visit Hamburg for 2 days.  
3364 \*\*Day 11:\*\* Fly from Hamburg to Porto.  
3365 \*\*Day 11-12:\*\* Visit Porto for 2 days.  
3366 \*\*Day 12:\*\* Fly from Porto to Frankfurt.  
3367 \*\*Day 12-13:\*\* Visit Frankfurt for 2 days.  
3368 \*\*Day 13:\*\* Fly from Frankfurt to Prague.  
3369 \*\*Day 13-14:\*\* Visit Prague for 2 days.  
3370 \*\*Day 14:\*\* Fly from Prague to Athens.  
3371 \*\*Day 14-17:\*\* Visit Athens for 4 days.  
3372 \*\*Day 17:\*\* Fly from Athens to Istanbul.  
3373 \*\*Day 17-21:\*\* Visit Istanbul for 5 days.  
3374 \*\*Day 21:\*\* Fly from Istanbul to Vilnius.  
3375 \*\*Day 21-22:\*\* Visit Vilnius for 2 days.

3376 TASK: You plan to visit 10 European cities for 23 days in total. You only take  
3377 direct flights to commute between cities. You would like to visit Stuttgart for  
3378 2 days. You would like to visit Split for 2 days. You are going to attend a  
3379 wedding in Split between day 22 and day  
3380 23. You would like to visit Vienna for 5 days. You want to spend 4 days in Madrid.  
3381 You plan to stay in Athens for 2 days. You would like to visit London for 3  
3382 days. During day 8 and day 10, you have to attend a conference in London. You  
3383 plan to stay in Paris for 3 days. You  
3384 want to meet a friend in Paris between day 10 and day 12. You plan to stay in  
3385 Reykjavik for 2 days. You have to attend a workshop in Reykjavik between day 16  
3386 and day 17. You want to spend 4 days in Seville. You want to spend 5 days in  
3387 Milan. You would like to meet your  
3388 friends at Milan between day 17 and day 21 to tour together.

3389 Here are the cities that have direct flights:  
3390 Athens and Paris, Athens and Vienna, Madrid and Vienna, Madrid and Split, Vienna and  
3391 Stuttgart, Paris and Milan, London and Vienna, London and Milan, Paris and  
3392 Reykjavik, Athens and London, from Reykjavik to Stuttgart, Seville and Milan,  
3393 from Reykjavik to Madrid, London and  
3394 Stuttgart, Milan and Stuttgart, Vienna and Reykjavik, Athens and Split, Athens and  
3395 Milan, Madrid and Athens, Madrid and London, Paris and Split, London and Paris,  
3396 Seville and Vienna, Vienna and Milan, Athens and Stuttgart, Madrid and Paris,  
3397 Seville and Madrid, from Reykjavik  
3398 to Athens, Vienna and Split, London and Split, Stuttgart and Split, Seville and  
3399 Paris, Paris and Stuttgart, Reykjavik and Milan, London and Reykjavik, Madrid  
3400 and Milan, Paris and Vienna, Milan and Split.

3401 Find a trip plan of visiting the cities for 23 days by taking direct flights to  
commute between them.

SOLUTION: Here is the trip plan for visiting the 10 European cities for 23 days:

\*\*Day 1-4:\*\* Arriving in Seville and visit Seville for 4 days.  
\*\*Day 4:\*\* Fly from Seville to Madrid.  
\*\*Day 4-7:\*\* Visit Madrid for 4 days.  
\*\*Day 7:\*\* Fly from Madrid to Athens.

3402 \*\*Day 7-8:\*\* Visit Athens for 2 days.  
3403 \*\*Day 8:\*\* Fly from Athens to London.  
3404 \*\*Day 8-10:\*\* Visit London for 3 days.  
3405 \*\*Day 10:\*\* Fly from London to Paris.  
3406 \*\*Day 10-12:\*\* Visit Paris for 3 days.  
3407 \*\*Day 12:\*\* Fly from Paris to Vienna.  
3408 \*\*Day 12-16:\*\* Visit Vienna for 5 days.  
3409 \*\*Day 16:\*\* Fly from Vienna to Reykjavik.  
3410 \*\*Day 16-17:\*\* Visit Reykjavik for 2 days.  
3411 \*\*Day 17:\*\* Fly from Reykjavik to Milan.  
3412 \*\*Day 17-21:\*\* Visit Milan for 5 days.  
3413 \*\*Day 21:\*\* Fly from Milan to Stuttgart.  
3414 \*\*Day 21-22:\*\* Visit Stuttgart for 2 days.  
3415 \*\*Day 22:\*\* Fly from Stuttgart to Split.  
3416 \*\*Day 22-23:\*\* Visit Split for 2 days.

3416 TASK: You plan to visit 10 European cities for 25 days in total. You only take  
3417 direct flights to commute between cities. You would like to visit Berlin for 2  
3418 days. You would like to visit Riga for 2 days. During day 5 and day 6, you have  
3419 to attend a conference in Riga. You  
3420 want to spend 3 days in Barcelona. You would like to visit Lyon for 4 days. You  
3421 would like to meet your friends at Lyon between day 8 and day 11 to tour  
3422 together. You plan to stay in Naples for 2 days. You plan to stay in Venice for  
3423 5 days. You want to spend 5 days in  
3424 Helsinki. You plan to visit relatives in Helsinki between day 21 and day 25. You  
3425 plan to stay in Rome for 5 days. You want to spend 3 days in Vilnius. You want  
3426 to spend 3 days in Amsterdam. You are going to attend a wedding in Amsterdam  
3427 between day 19 and day 21.

3427 Here are the cities that have direct flights:  
3428 Berlin and Amsterdam, Rome and Helsinki, Rome and Lyon, Naples and Amsterdam, Riga  
3429 and Barcelona, Rome and Venice, Riga and Amsterdam, from Riga to Vilnius,  
3430 Barcelona and Berlin, Rome and Barcelona, Rome and Amsterdam, Barcelona and  
3431 Venice, Berlin and Helsinki, Amsterdam and  
3432 Helsinki, Vilnius and Helsinki, Rome and Berlin, from Rome to Riga, Barcelona and  
3433 Amsterdam, Venice and Naples, Barcelona and Lyon, Naples and Berlin, Barcelona  
3434 and Helsinki, Venice and Helsinki, Barcelona and Naples, Vilnius and Amsterdam,  
3435 Venice and Amsterdam, Lyon and  
3436 Venice, Naples and Helsinki, Riga and Berlin, Rome and Naples, Venice and Berlin,  
3437 Riga and Helsinki, Berlin and Vilnius, Lyon and Amsterdam.

3436 Find a trip plan of visiting the cities for 25 days by taking direct flights to  
3437 commute between them.

3438 SOLUTION: Here is the trip plan for visiting the 10 European cities for 25 days:  
3439

3440 \*\*Day 1-5:\*\* Arriving in Rome and visit Rome for 5 days.  
3441 \*\*Day 5:\*\* Fly from Rome to Riga.  
3442 \*\*Day 5-6:\*\* Visit Riga for 2 days.  
3443 \*\*Day 6:\*\* Fly from Riga to Barcelona.  
3444 \*\*Day 6-8:\*\* Visit Barcelona for 3 days.  
3445 \*\*Day 8:\*\* Fly from Barcelona to Lyon.  
3446 \*\*Day 8-11:\*\* Visit Lyon for 4 days.  
3447 \*\*Day 11:\*\* Fly from Lyon to Venice.  
3448 \*\*Day 11-15:\*\* Visit Venice for 5 days.  
3449 \*\*Day 15:\*\* Fly from Venice to Naples.  
3450 \*\*Day 15-16:\*\* Visit Naples for 2 days.  
3451 \*\*Day 16:\*\* Fly from Naples to Berlin.  
3452 \*\*Day 16-17:\*\* Visit Berlin for 2 days.  
3453 \*\*Day 17:\*\* Fly from Berlin to Vilnius.  
3454 \*\*Day 17-19:\*\* Visit Vilnius for 3 days.  
3455 \*\*Day 19:\*\* Fly from Vilnius to Amsterdam.  
3456 \*\*Day 19-21:\*\* Visit Amsterdam for 3 days.  
3457 \*\*Day 21:\*\* Fly from Amsterdam to Helsinki.  
3458 \*\*Day 21-25:\*\* Visit Helsinki for 5 days.



3456 TASK: You plan to visit 10 European cities for 27 days in total. You only take  
3457 direct flights to commute between cities. You would like to visit Prague for 5  
3458 days. You have to attend a workshop in Prague between day 7 and day 11. You  
3459 would like to visit Helsinki for 3 days.  
3460 You are going to attend a wedding in Helsinki between day 15 and day 17. You plan to  
3461 stay in Tallinn for 2 days. You want to meet a friend in Tallinn between day 6  
3462 and day 7. You want to spend 4 days in Edinburgh. You want to spend 5 days in  
3463 Paris. You want to spend 4 days  
3464 in Vienna. You plan to stay in Lisbon for 5 days. From day 11 to day 15, there is a  
3465 annual show you want to attend in Lisbon. You want to spend 4 days in Budapest.  
3466 You plan to stay in Lyon for 2 days. You plan to stay in Brussels for 2 days.  
3467 You would like to meet your  
3468 friends at Brussels between day 1 and day 2 to tour together.

3468 Here are the cities that have direct flights:  
3469 Prague and Lyon, Brussels and Lisbon, Helsinki and Budapest, Vienna and Lyon, Paris  
3470 and Tallinn, Brussels and Prague, Brussels and Helsinki, Prague and Helsinki,  
3471 Brussels and Vienna, Brussels and Budapest, Lisbon and Budapest, Tallinn and  
3472 Helsinki, Brussels and Paris,  
3473 Brussels and Tallinn, Lisbon and Lyon, Prague and Lisbon, Paris and Prague, Helsinki  
3474 and Edinburgh, Prague and Edinburgh, Tallinn and Prague, Brussels and Lyon,  
3475 Paris and Lisbon, Helsinki and Vienna, Paris and Helsinki, Paris and Budapest,  
3476 Edinburgh and Budapest, Brussels  
3477 and Edinburgh, Lisbon and Vienna, Paris and Lyon, Lisbon and Helsinki, Prague and  
3478 Vienna, Paris and Vienna, Prague and Budapest, Paris and Edinburgh, Budapest  
3479 and Vienna.

3479 Find a trip plan of visiting the cities for 27 days by taking direct flights to  
3480 commute between them.  
3481 SOLUTION: Here is the trip plan for visiting the 10 European cities for 27 days:

3482 \*\*Day 1-2:\*\* Arriving in Brussels and visit Brussels for 2 days.  
3483 \*\*Day 2:\*\* Fly from Brussels to Paris.  
3484 \*\*Day 2-6:\*\* Visit Paris for 5 days.  
3485 \*\*Day 6:\*\* Fly from Paris to Tallinn.  
3486 \*\*Day 6-7:\*\* Visit Tallinn for 2 days.  
3487 \*\*Day 7:\*\* Fly from Tallinn to Prague.  
3488 \*\*Day 7-11:\*\* Visit Prague for 5 days.  
3489 \*\*Day 11:\*\* Fly from Prague to Lisbon.  
3490 \*\*Day 11-15:\*\* Visit Lisbon for 5 days.  
3491 \*\*Day 15:\*\* Fly from Lisbon to Helsinki.  
3492 \*\*Day 15-17:\*\* Visit Helsinki for 3 days.  
3493 \*\*Day 17:\*\* Fly from Helsinki to Edinburgh.  
3494 \*\*Day 17-20:\*\* Visit Edinburgh for 4 days.  
3495 \*\*Day 20:\*\* Fly from Edinburgh to Budapest.  
3496 \*\*Day 20-23:\*\* Visit Budapest for 4 days.  
3497 \*\*Day 23:\*\* Fly from Budapest to Vienna.  
3498 \*\*Day 23-26:\*\* Visit Vienna for 4 days.  
3499 \*\*Day 26:\*\* Fly from Vienna to Lyon.  
3500 \*\*Day 26-27:\*\* Visit Lyon for 2 days.

3499 Query:  
3500 You plan to visit 10 European cities for 25 days in total. You only take direct  
3501 flights to commute between cities. You plan to stay in Berlin for 4 days. You  
3502 have to attend a workshop in Berlin between day 12 and day 15. You would like  
3503 to visit Prague for 2 days. You plan to  
3504 stay in Stuttgart for 5 days. You would like to visit Manchester for 3 days. You  
3505 want to spend 2 days in Nice. You would like to visit Reykjavik for 2 days. You  
3506 would like to visit Florence for 3 days. You plan to stay in Vilnius for 5  
3507 days. You would like to meet your  
3508 friends at Vilnius between day 15 and day 19 to tour together. You plan to stay in  
3509 Oslo for 4 days. You would like to visit Dubrovnik for 4 days. You plan to  
visit relatives in Dubrovnik between day 1 and day 4.

---

3510 Here are the cities that have direct flights:  
3511 from Reykjavik to Stuttgart, Manchester and Stuttgart, Nice and Berlin, Oslo and  
3512 Prague, Stuttgart and Berlin, Manchester and Nice, Reykjavik and Oslo,  
3513 Reykjavik and Prague, Manchester and Prague, Reykjavik and Berlin, Dubrovnik  
3514 and Manchester, Manchester and Oslo,  
3515 Manchester and Berlin, Prague and Florence, Berlin and Vilnius, Dubrovnik and Oslo,  
3516 Nice and Oslo, Berlin and Oslo, Nice and Reykjavik, Vilnius and Oslo.

3517 Find a trip plan of visiting the cities for 25 days by taking direct flights to  
3518 commute between them.

3519

3520 =====RESPONSE OF o1-preview=====

3521 SOLUTION: Here is the trip plan for visiting the 10 European cities for 25 days:  
3522

3523 \*\*Day 1-4:\*\* Arriving in Dubrovnik and visit Dubrovnik for 4 days.  
3524  
3525 \*\*Day 4:\*\* Fly from Dubrovnik to Manchester.  
3526  
3527 \*\*Day 4-7:\*\* Visit Manchester for 3 days.  
3528  
3529 \*\*Day 7:\*\* Fly from Manchester to Stuttgart.  
3530  
3531 \*\*Day 7-12:\*\* Visit Stuttgart for 5 days.  
3532  
3533 \*\*Day 12:\*\* Fly from Stuttgart to Berlin.  
3534  
3535 \*\*Day 12-15:\*\* Visit Berlin for 4 days. (Attend workshop between day 12 and day 15)  
3536  
3537 \*\*Day 15:\*\* Fly from Berlin to Vilnius.  
3538  
3539 \*\*Day 15-19:\*\* Visit Vilnius for 5 days. (Meet friends between day 15 and day 19)  
3540  
3541 \*\*Day 19:\*\* Fly from Vilnius to Oslo via Berlin.  
3542  
3543 \*\*Day 19-23:\*\* Visit Oslo for 4 days.  
3544  
3545 \*\*Day 23:\*\* Fly from Oslo to Reykjavik.  
3546  
3547 \*\*Day 23-25:\*\* Visit Reykjavik for 2 days.

3548 =====FEEDBACK FROM METACONTROLLER=====

3549 Number of cities in plan is 7, expected 10

3550

3551 P.2 BACK PROMPT

3552 You are an expert at planning trips. Fix the below given trip schedule such that it  
3553 meets as many specified constraints as possible. You are given a few  
3554 constraints regarding the cities to visit and the durations of staying at each  
3555 city. You are also given the flight  
3556 information between the cities.  
3557 Produce only one plan, strictly adhering to the format shown in the examples below,  
3558 without any extra information.

3559

3560 The produced plan will be parsed using the following python code:

3561

```
3562 '''python
3563 def parse_response(response: str):
3564     """Parse the response.
```

3565

3566 Returns a parsed plan in a list of (city, stay\_days) tuples.

3567

3568 Args:

3569 response: Raw response from the model.

---

```

3564
3565 Returns:
3566     Structured plan after parsing.
3567     """
3568     pattern_visit = r'\d+-\d+'
3569     pattern_flight = r'.*Day (\d+).*from (\w+) to (\w+)'
3570     pattern_days = r'European cities for (\d+) days'
3571
3572     days, flights, flight_days = [], [], []
3573     total_days = None
3574     for piece in response.split('\n'):
3575         days_match = re.findall(pattern_days, piece)
3576         if days_match:
3577             total_days = int(days_match[0])
3578
3579         visit_match = re.findall(pattern_visit, piece)
3580         if visit_match:
3581             days.append(visit_match[0])
3582             end_day = int(visit_match[0].split('-')[1])
3583             # Reach the end of the plan, stop to avoid parsing alternative plans.
3584             if end_day == total_days:
3585                 break
3586
3587         flight_match = re.findall(pattern_flight, piece)
3588         if flight_match:
3589             flights.append(flight_match[0])
3590
3591     visit_cities, parsed_plan = [], []
3592     for flight_day, begin_city, end_city in flights:
3593         flight_days.append(int(flight_day))
3594         if not visit_cities:
3595             visit_cities.append(begin_city)
3596             visit_cities.append(end_city)
3597         else:
3598             visit_cities.append(end_city)
3599
3600     if not days or not flights or not visit_cities:
3601         return []
3602     last_day = int(days[-1].split('-')[1])
3603     flight_days = [1] + flight_days +
3604     for i, visit_city in enumerate(visit_cities):
3605         city_stay = flight_days - flight_days + 1
3606         parsed_plan.append((visit_city, city_stay))
3607
3608     return parsed_plan
3609 '''
3610
3611 Here are a few example tasks and solutions:
3612
3613 TASK: You plan to visit 10 European cities for 21 days in total. You only take
3614 direct flights to commute between cities. You plan to stay in Vienna for 3 days.
3615 You want to spend 5 days in Frankfurt. You want to spend 2 days in Oslo. You
3616 are going to attend a wedding in Oslo
3617 between day 20 and day 21. You want to spend 3 days in Prague. You would like to
3618 visit Valencia for 2 days. You want to meet a friend in Valencia between day 17
3619 and day 18. You plan to stay in Dubrovnik for 2 days. You would like to visit
3620 Edinburgh for 5 days. From day 1 to
3621 day 5, there is a annual show you want to attend in Edinburgh. You plan to stay in
3622 London for 2 days. You plan to visit relatives in London between day 12 and day
3623 13. You would like to visit Munich for 3 days. You would like to meet your
3624 friends at Munich between day 18 and
3625 day 20 to tour together. You would like to visit Budapest for 3 days.
3626
3627 Here are the cities that have direct flights:
3628 Valencia and Munich, Vienna and Munich, Vienna and Valencia, London and Budapest,
3629 London and Oslo, Edinburgh and Budapest, Frankfurt and Budapest, Frankfurt and

```

3618 London, Prague and Oslo, Edinburgh and Oslo, Edinburgh and Munich, Prague and  
3619 Munich, London and Prague, Edinburgh  
3620 and London, Edinburgh and Frankfurt, Dubrovnik and Munich, Dubrovnik and Vienna,  
3621 Munich and Oslo, Dubrovnik and Oslo, Budapest and Munich, Frankfurt and Prague,  
3622 Vienna and London, Frankfurt and Vienna, Frankfurt and Oslo, Frankfurt and  
3623 Munich, Vienna and Oslo, Vienna and  
3624 Prague, Budapest and Oslo, Budapest and Prague, London and Valencia, London and  
3625 Munich, Frankfurt and Dubrovnik, Prague and Valencia, Frankfurt and Valencia,  
3626 Edinburgh and Prague, Vienna and Budapest.

3627 Find a trip plan of visiting the cities for 21 days by taking direct flights to  
3628 commute between them.

3629 SOLUTION: Here is the trip plan for visiting the 10 European cities for 21 days:

3630 \*\*Day 1-5:\*\* Arriving in Edinburgh and visit Edinburgh for 5 days.  
3631 \*\*Day 5:\*\* Fly from Edinburgh to Frankfurt.  
3632 \*\*Day 5-9:\*\* Visit Frankfurt for 5 days.  
3633 \*\*Day 9:\*\* Fly from Frankfurt to Dubrovnik.  
3634 \*\*Day 9-10:\*\* Visit Dubrovnik for 2 days.  
3635 \*\*Day 10:\*\* Fly from Dubrovnik to Vienna.  
3636 \*\*Day 10-12:\*\* Visit Vienna for 3 days.  
3637 \*\*Day 12:\*\* Fly from Vienna to London.  
3638 \*\*Day 12-13:\*\* Visit London for 2 days.  
3639 \*\*Day 13:\*\* Fly from London to Budapest.  
3640 \*\*Day 13-15:\*\* Visit Budapest for 3 days.  
3641 \*\*Day 15:\*\* Fly from Budapest to Prague.  
3642 \*\*Day 15-17:\*\* Visit Prague for 3 days.  
3643 \*\*Day 17:\*\* Fly from Prague to Valencia.  
3644 \*\*Day 17-18:\*\* Visit Valencia for 2 days.  
3645 \*\*Day 18:\*\* Fly from Valencia to Munich.  
3646 \*\*Day 18-20:\*\* Visit Munich for 3 days.  
3647 \*\*Day 20:\*\* Fly from Munich to Oslo.  
3648 \*\*Day 20-21:\*\* Visit Oslo for 2 days.

3649 TASK: You plan to visit 10 European cities for 22 days in total. You only take  
3650 direct flights to commute between cities. You want to spend 4 days in Helsinki.  
3651 You plan to stay in Vilnius for 2 days. You plan to stay in Istanbul for 5  
3652 days. From day 17 to day 21, there is a  
3653 annual show you want to attend in Istanbul. You want to spend 2 days in Hamburg. You  
3654 plan to stay in Porto for 2 days. During day 11 and day 12, you have to attend  
3655 a conference in Porto. You want to spend 4 days in Athens. You would like to  
3656 visit Prague for 2 days. You plan  
3657 to visit relatives in Prague between day 13 and day 14. You want to spend 2 days in  
3658 Frankfurt. You want to spend 3 days in Krakow. You want to spend 5 days in  
3659 Munich. You would like to meet your friends at Munich between day 1 and day 5  
3660 to tour together.

3661 Here are the cities that have direct flights:  
3662 from Krakow to Vilnius, Helsinki and Hamburg, Hamburg and Athens, Munich and  
3663 Frankfurt, Hamburg and Porto, Munich and Istanbul, Prague and Athens, Frankfurt  
3664 and Athens, Munich and Athens, Munich and Prague, from Vilnius to Munich,  
3665 Hamburg and Istanbul, Frankfurt and  
3666 Istanbul, Munich and Krakow, Munich and Hamburg, Munich and Helsinki, Prague and  
3667 Istanbul, Frankfurt and Vilnius, Helsinki and Istanbul, Athens and Vilnius,  
3668 Krakow and Frankfurt, Helsinki and Frankfurt, Porto and Frankfurt, Frankfurt  
3669 and Prague, Istanbul and Vilnius, Krakow  
3670 and Istanbul, Krakow and Prague, Munich and Porto, Helsinki and Vilnius, Helsinki  
3671 and Prague, Porto and Istanbul, Hamburg and Frankfurt, Krakow and Helsinki,  
Athens and Istanbul.

3672 Find a trip plan of visiting the cities for 22 days by taking direct flights to  
3673 commute between them.

3674 SOLUTION: Here is the trip plan for visiting the 10 European cities for 22 days:

3675 \*\*Day 1-5:\*\* Arriving in Munich and visit Munich for 5 days.

3672 \*\*Day 5:\*\* Fly from Munich to Krakow.  
3673 \*\*Day 5-7:\*\* Visit Krakow for 3 days.  
3674 \*\*Day 7:\*\* Fly from Krakow to Helsinki.  
3675 \*\*Day 7-10:\*\* Visit Helsinki for 4 days.  
3676 \*\*Day 10:\*\* Fly from Helsinki to Hamburg.  
3677 \*\*Day 10-11:\*\* Visit Hamburg for 2 days.  
3678 \*\*Day 11:\*\* Fly from Hamburg to Porto.  
3679 \*\*Day 11-12:\*\* Visit Porto for 2 days.  
3680 \*\*Day 12:\*\* Fly from Porto to Frankfurt.  
3681 \*\*Day 12-13:\*\* Visit Frankfurt for 2 days.  
3682 \*\*Day 13:\*\* Fly from Frankfurt to Prague.  
3683 \*\*Day 13-14:\*\* Visit Prague for 2 days.  
3684 \*\*Day 14:\*\* Fly from Prague to Athens.  
3685 \*\*Day 14-17:\*\* Visit Athens for 4 days.  
3686 \*\*Day 17:\*\* Fly from Athens to Istanbul.  
3687 \*\*Day 17-21:\*\* Visit Istanbul for 5 days.  
3688 \*\*Day 21:\*\* Fly from Istanbul to Vilnius.  
3689 \*\*Day 21-22:\*\* Visit Vilnius for 2 days.

3688 TASK: You plan to visit 10 European cities for 23 days in total. You only take  
3689 direct flights to commute between cities. You would like to visit Stuttgart for  
3690 2 days. You would like to visit Split for 2 days. You are going to attend a  
3691 wedding in Split between day 22 and day  
3692 23. You would like to visit Vienna for 5 days. You want to spend 4 days in Madrid.  
3693 You plan to stay in Athens for 2 days. You would like to visit London for 3  
3694 days. During day 8 and day 10, you have to attend a conference in London. You  
3695 plan to stay in Paris for 3 days. You  
3696 want to meet a friend in Paris between day 10 and day 12. You plan to stay in  
3697 Reykjavik for 2 days. You have to attend a workshop in Reykjavik between day 16  
3698 and day 17. You want to spend 4 days in Seville. You want to spend 5 days in  
3699 Milan. You would like to meet your  
3700 friends at Milan between day 17 and day 21 to tour together.

3700 Here are the cities that have direct flights:  
3701 Athens and Paris, Athens and Vienna, Madrid and Vienna, Madrid and Split, Vienna and  
3702 Stuttgart, Paris and Milan, London and Vienna, London and Milan, Paris and  
3703 Reykjavik, Athens and London, from Reykjavik to Stuttgart, Seville and Milan,  
3704 from Reykjavik to Madrid, London and  
3705 Stuttgart, Milan and Stuttgart, Vienna and Reykjavik, Athens and Split, Athens and  
3706 Milan, Madrid and Athens, Madrid and London, Paris and Split, London and Paris,  
3707 Seville and Vienna, Vienna and Milan, Athens and Stuttgart, Madrid and Paris,  
3708 Seville and Madrid, from Reykjavik  
3709 to Athens, Vienna and Split, London and Split, Stuttgart and Split, Seville and  
3710 Paris, Paris and Stuttgart, Reykjavik and Milan, London and Reykjavik, Madrid  
3711 and Milan, Paris and Vienna, Milan and Split.

3711 Find a trip plan of visiting the cities for 23 days by taking direct flights to  
3712 commute between them.  
3713 SOLUTION: Here is the trip plan for visiting the 10 European cities for 23 days:

3714 \*\*Day 1-4:\*\* Arriving in Seville and visit Seville for 4 days.  
3715 \*\*Day 4:\*\* Fly from Seville to Madrid.  
3716 \*\*Day 4-7:\*\* Visit Madrid for 4 days.  
3717 \*\*Day 7:\*\* Fly from Madrid to Athens.  
3718 \*\*Day 7-8:\*\* Visit Athens for 2 days.  
3719 \*\*Day 8:\*\* Fly from Athens to London.  
3720 \*\*Day 8-10:\*\* Visit London for 3 days.  
3721 \*\*Day 10:\*\* Fly from London to Paris.  
3722 \*\*Day 10-12:\*\* Visit Paris for 3 days.  
3723 \*\*Day 12:\*\* Fly from Paris to Vienna.  
3724 \*\*Day 12-16:\*\* Visit Vienna for 5 days.  
3725 \*\*Day 16:\*\* Fly from Vienna to Reykjavik.  
3726 \*\*Day 16-17:\*\* Visit Reykjavik for 2 days.  
3727 \*\*Day 17:\*\* Fly from Reykjavik to Milan.  
3728 \*\*Day 17-21:\*\* Visit Milan for 5 days.

3726 \*\*Day 21:\*\* Fly from Milan to Stuttgart.  
3727 \*\*Day 21-22:\*\* Visit Stuttgart for 2 days.  
3728 \*\*Day 22:\*\* Fly from Stuttgart to Split.  
3729 \*\*Day 22-23:\*\* Visit Split for 2 days.  
3730  
3731 TASK: You plan to visit 10 European cities for 25 days in total. You only take  
3732 direct flights to commute between cities. You would like to visit Berlin for 2  
3733 days. You would like to visit Riga for 2 days. During day 5 and day 6, you have  
3734 to attend a conference in Riga. You  
3735 want to spend 3 days in Barcelona. You would like to visit Lyon for 4 days. You  
3736 would like to meet your friends at Lyon between day 8 and day 11 to tour  
3737 together. You plan to stay in Naples for 2 days. You plan to stay in Venice for  
3738 5 days. You want to spend 5 days in  
3739 Helsinki. You plan to visit relatives in Helsinki between day 21 and day 25. You  
3740 plan to stay in Rome for 5 days. You want to spend 3 days in Vilnius. You want  
3741 to spend 3 days in Amsterdam. You are going to attend a wedding in Amsterdam  
3742 between day 19 and day 21.  
3743  
3744 Here are the cities that have direct flights:  
3745 Berlin and Amsterdam, Rome and Helsinki, Rome and Lyon, Naples and Amsterdam, Riga  
3746 and Barcelona, Rome and Venice, Riga and Amsterdam, from Riga to Vilnius,  
3747 Barcelona and Berlin, Rome and Barcelona, Rome and Amsterdam, Barcelona and  
3748 Venice, Berlin and Helsinki, Amsterdam and  
3749 Helsinki, Vilnius and Helsinki, Rome and Berlin, from Rome to Riga, Barcelona and  
3750 Amsterdam, Venice and Naples, Barcelona and Lyon, Naples and Berlin, Barcelona  
3751 and Helsinki, Venice and Helsinki, Barcelona and Naples, Vilnius and Amsterdam,  
3752 Venice and Amsterdam, Lyon and  
3753 Venice, Naples and Helsinki, Riga and Berlin, Rome and Naples, Venice and Berlin,  
3754 Riga and Helsinki, Berlin and Vilnius, Lyon and Amsterdam.  
3755  
3756 Find a trip plan of visiting the cities for 25 days by taking direct flights to  
3757 commute between them.  
3758 SOLUTION: Here is the trip plan for visiting the 10 European cities for 25 days:  
3759  
3760 \*\*Day 1-5:\*\* Arriving in Rome and visit Rome for 5 days.  
3761 \*\*Day 5:\*\* Fly from Rome to Riga.  
3762 \*\*Day 5-6:\*\* Visit Riga for 2 days.  
3763 \*\*Day 6:\*\* Fly from Riga to Barcelona.  
3764 \*\*Day 6-8:\*\* Visit Barcelona for 3 days.  
3765 \*\*Day 8:\*\* Fly from Barcelona to Lyon.  
3766 \*\*Day 8-11:\*\* Visit Lyon for 4 days.  
3767 \*\*Day 11:\*\* Fly from Lyon to Venice.  
3768 \*\*Day 11-15:\*\* Visit Venice for 5 days.  
3769 \*\*Day 15:\*\* Fly from Venice to Naples.  
3770 \*\*Day 15-16:\*\* Visit Naples for 2 days.  
3771 \*\*Day 16:\*\* Fly from Naples to Berlin.  
3772 \*\*Day 16-17:\*\* Visit Berlin for 2 days.  
3773 \*\*Day 17:\*\* Fly from Berlin to Vilnius.  
3774 \*\*Day 17-19:\*\* Visit Vilnius for 3 days.  
3775 \*\*Day 19:\*\* Fly from Vilnius to Amsterdam.  
3776 \*\*Day 19-21:\*\* Visit Amsterdam for 3 days.  
3777 \*\*Day 21:\*\* Fly from Amsterdam to Helsinki.  
3778 \*\*Day 21-25:\*\* Visit Helsinki for 5 days.  
3779  
3780 TASK: You plan to visit 10 European cities for 27 days in total. You only take  
3781 direct flights to commute between cities. You would like to visit Prague for 5  
3782 days. You have to attend a workshop in Prague between day 7 and day 11. You  
3783 would like to visit Helsinki for 3 days.  
3784 You are going to attend a wedding in Helsinki between day 15 and day 17. You plan to  
3785 stay in Tallinn for 2 days. You want to meet a friend in Tallinn between day 6  
3786 and day 7. You want to spend 4 days in Edinburgh. You want to spend 5 days in  
3787 Paris. You want to spend 4 days  
3788 in Vienna. You plan to stay in Lisbon for 5 days. From day 11 to day 15, there is a  
3789 annual show you want to attend in Lisbon. You want to spend 4 days in Budapest.

3780           You plan to stay in Lyon for 2 days. You plan to stay in Brussels for 2 days.  
3781           You would like to meet your  
3782 friends at Brussels between day 1 and day 2 to tour together.  
3783  
3784 Here are the cities that have direct flights:  
3785 Prague and Lyon, Brussels and Lisbon, Helsinki and Budapest, Vienna and Lyon, Paris  
3786           and Tallinn, Brussels and Prague, Brussels and Helsinki, Prague and Helsinki,  
3787           Brussels and Vienna, Brussels and Budapest, Lisbon and Budapest, Tallinn and  
3788           Helsinki, Brussels and Paris,  
3789 Brussels and Tallinn, Lisbon and Lyon, Prague and Lisbon, Paris and Prague, Helsinki  
3790           and Edinburgh, Prague and Edinburgh, Tallinn and Prague, Brussels and Lyon,  
3791           Paris and Lisbon, Helsinki and Vienna, Paris and Helsinki, Paris and Budapest,  
3792           Edinburgh and Budapest, Brussels  
3793           and Edinburgh, Lisbon and Vienna, Paris and Lyon, Lisbon and Helsinki, Prague and  
3794           Vienna, Paris and Vienna, Prague and Budapest, Paris and Edinburgh, Budapest  
3795           and Vienna.  
3796 Find a trip plan of visiting the cities for 27 days by taking direct flights to  
3797 commute between them.  
3798 SOLUTION: Here is the trip plan for visiting the 10 European cities for 27 days:  
3799  
3800 \*\*Day 1-2:\*\* Arriving in Brussels and visit Brussels for 2 days.  
3801 \*\*Day 2:\*\* Fly from Brussels to Paris.  
3802 \*\*Day 2-6:\*\* Visit Paris for 5 days.  
3803 \*\*Day 6:\*\* Fly from Paris to Tallinn.  
3804 \*\*Day 6-7:\*\* Visit Tallinn for 2 days.  
3805 \*\*Day 7:\*\* Fly from Tallinn to Prague.  
3806 \*\*Day 7-11:\*\* Visit Prague for 5 days.  
3807 \*\*Day 11:\*\* Fly from Prague to Lisbon.  
3808 \*\*Day 11-15:\*\* Visit Lisbon for 5 days.  
3809 \*\*Day 15:\*\* Fly from Lisbon to Helsinki.  
3810 \*\*Day 15-17:\*\* Visit Helsinki for 3 days.  
3811 \*\*Day 17:\*\* Fly from Helsinki to Edinburgh.  
3812 \*\*Day 17-20:\*\* Visit Edinburgh for 4 days.  
3813 \*\*Day 20:\*\* Fly from Edinburgh to Budapest.  
3814 \*\*Day 20-23:\*\* Visit Budapest for 4 days.  
3815 \*\*Day 23:\*\* Fly from Budapest to Vienna.  
3816 \*\*Day 23-26:\*\* Visit Vienna for 4 days.  
3817 \*\*Day 26:\*\* Fly from Vienna to Lyon.  
3818 \*\*Day 26-27:\*\* Visit Lyon for 2 days.  
3819  
3820 Query:  
3821 You plan to visit 10 European cities for 25 days in total. You only take direct  
3822 flights to commute between cities. You plan to stay in Berlin for 4 days. You  
3823 have to attend a workshop in Berlin between day 12 and day 15. You would like  
3824 to visit Prague for 2 days. You plan to  
3825 stay in Stuttgart for 5 days. You would like to visit Manchester for 3 days. You  
3826 want to spend 2 days in Nice. You would like to visit Reykjavik for 2 days. You  
3827 would like to visit Florence for 3 days. You plan to stay in Vilnius for 5  
3828 days. You would like to meet your  
3829 friends at Vilnius between day 15 and day 19 to tour together. You plan to stay in  
3830 Oslo for 4 days. You would like to visit Dubrovnik for 4 days. You plan to  
3831 visit relatives in Dubrovnik between day 1 and day 4.  
3832  
3833 Here are the cities that have direct flights:  
3834 from Reykjavik to Stuttgart, Manchester and Stuttgart, Nice and Berlin, Oslo and  
3835 Prague, Stuttgart and Berlin, Manchester and Nice, Reykjavik and Oslo,  
3836 Reykjavik and Prague, Manchester and Prague, Reykjavik and Berlin, Dubrovnik  
3837 and Manchester, Manchester and Oslo,  
3838 Manchester and Berlin, Prague and Florence, Berlin and Vilnius, Dubrovnik and Oslo,  
3839 Nice and Oslo, Berlin and Oslo, Nice and Reykjavik, Vilnius and Oslo.  
3840 Find a trip plan of visiting the cities for 25 days by taking direct flights to  
3841 commute between them.

---

3834  
3835 Incorrect plan in natural language:  
3836 SOLUTION: Here is the trip plan for visiting the 10 European cities for 25 days:  
3837  
3838 \*\*Day 1-4:\*\* Arriving in Dubrovnik and visit Dubrovnik for 4 days.  
3839  
3840 \*\*Day 4:\*\* Fly from Dubrovnik to Manchester.  
3841  
3842 \*\*Day 4-7:\*\* Visit Manchester for 3 days.  
3843  
3844 \*\*Day 7:\*\* Fly from Manchester to Stuttgart.  
3845  
3846 \*\*Day 7-12:\*\* Visit Stuttgart for 5 days.  
3847  
3848 \*\*Day 12:\*\* Fly from Stuttgart to Berlin.  
3849  
3850 \*\*Day 12-15:\*\* Visit Berlin for 4 days. (Attend workshop between day 12 and day 15)  
3851  
3852 \*\*Day 15:\*\* Fly from Berlin to Vilnius.  
3853  
3854 \*\*Day 15-19:\*\* Visit Vilnius for 5 days. (Meet friends between day 15 and day 19)  
3855  
3856 \*\*Day 19:\*\* Fly from Vilnius to Oslo via Berlin.  
3857  
3858 \*\*Day 19-23:\*\* Visit Oslo for 4 days.  
3859  
3860 \*\*Day 23:\*\* Fly from Oslo to Reykjavik.  
3861  
3862 \*\*Day 23-25:\*\* Visit Reykjavik for 2 days.  
3863  
3864 Incorrect plan in parsed format:  
3865 [[ 'Dubrovnik', 4], [ 'Manchester', 4], [ 'Stuttgart', 6], [ 'Berlin', 4], [ 'Vilnius',  
3866 5], [ 'Oslo', 5], [ 'Reykjavik', 3]]  
3867  
3868 Errors with the above plan:  
3869 Number of cities in plan is 7, expected 10  
3870  
3871 Give the corrected plan in natural language such that it can be parsed by the above  
3872 python code.  
3873 Strictly adhere to the format shown in the examples below, without any extra  
3874 information.  
3875  
3876 Corrected plan:  
3877  
3878 =====RESPONSE OF o1-preview=====  
3879 SOLUTION: Here is the trip plan for visiting the 10 European cities for 25 days:  
3880  
3881 \*\*Day 1-4:\*\* Arriving in Dubrovnik and visit Dubrovnik for 4 days. (Visit relatives  
3882 between day 1 and day 4)  
3883  
3884 \*\*Day 4:\*\* Fly from Dubrovnik to Manchester.  
3885  
3886 \*\*Day 4-6:\*\* Visit Manchester for 3 days.  
3887  
3888 \*\*Day 6:\*\* Fly from Manchester to Nice.  
3889  
3890 \*\*Day 6-7:\*\* Visit Nice for 2 days.  
3891  
3892 \*\*Day 7:\*\* Fly from Nice to Reykjavik.  
3893  
3894 \*\*Day 7-8:\*\* Visit Reykjavik for 2 days.  
3895  
3896 \*\*Day 8:\*\* Fly from Reykjavik to Stuttgart.  
3897  
3898 \*\*Day 8-10:\*\* Visit Stuttgart for 3 days.



---

3888  
3889 \*\*Day 10:\*\* Fly from Stuttgart to Prague.  
3890  
3891 \*\*Day 10-11:\*\* Visit Prague for 2 days.  
3892  
3893 \*\*Day 11:\*\* Fly from Prague to Berlin.  
3894  
3895 \*\*Day 11-15:\*\* Visit Berlin for 4 days. (Attend workshop between day 12 and day 15)  
3896  
3897 \*\*Day 15:\*\* Fly from Berlin to Vilnius.  
3898  
3899 \*\*Day 15-19:\*\* Visit Vilnius for 5 days. (Meet friends between day 15 and day 19)  
3900  
3901 \*\*Day 19-22:\*\* Visit Oslo for 4 days.  
3902  
3903 \*\*Day 22:\*\* Fly from Oslo to Florence via Prague.  
3904  
3905 \*\*Day 22-25:\*\* Visit Florence for 3 days.  
3906  
3907 Q PROMPTS FOR CALENDAR SCHEDULING  
3908 Q.1 FIRST ITERATION  
3909  
3910 You are an expert at scheduling meetings. You are given a few constraints on the  
3911 existing schedule of each participant, the meeting duration, and possibly some  
3912 preferences on the meeting time. Note there exists a solution that works with  
3913 existing schedule of every  
3914 participant. Produce only one meeting time, strictly adhering to the format shown in  
3915 the examples below, without providing any extra information. Here are a few  
3916 example tasks and solutions:  
3917  
3918 TASK: You need to schedule a meeting for George, Steven, Aaron, Patrick and Cynthia  
3919 for half an hour between the work hours of 9:00 to 17:00 on Monday.  
3920  
3921 Here are the existing schedules for everyone during the day:  
3922 George has no meetings the whole day.  
3923 Steven is free the entire day.  
3924 Aaron has blocked their calendar on Monday during 9:00 to 10:00, 11:30 to 12:00,  
3925 15:30 to 17:00;  
3926 Patrick has blocked their calendar on Monday during 9:00 to 9:30, 10:00 to 11:00,  
3927 11:30 to 12:00, 12:30 to 14:00, 15:00 to 15:30;  
3928 Cynthia is busy on Monday during 9:00 to 9:30, 10:30 to 11:30, 12:30 to 14:30, 15:00  
3929 to 16:30;  
3930  
3931 Aaron can not meet on Monday after 12:30. Find a time that works for everyone's  
3932 schedule and constraints.  
3933 SOLUTION: Here is the proposed time: Monday, 12:00 - 12:30  
3934  
3935 TASK: You need to schedule a meeting for Elizabeth, Eugene, Nancy, Justin and Roy  
3936 for half an hour between the work hours of 9:00 to 17:00 on Monday.  
3937  
3938 Here are the existing schedules for everyone during the day:  
3939 Elizabeth's calendar is wide open the entire day.  
3940 Eugene has blocked their calendar on Monday during 12:00 to 12:30, 13:30 to 14:00,  
3941 15:00 to 16:00;  
3942 Nancy has meetings on Monday during 10:30 to 11:00, 12:00 to 13:00, 14:00 to 14:30,  
3943 15:00 to 16:00;  
3944 Justin has meetings on Monday during 9:30 to 10:00, 10:30 to 11:30, 12:00 to 12:30,  
3945 13:00 to 13:30, 14:30 to 17:00;  
3946 Roy is busy on Monday during 10:00 to 11:00, 11:30 to 12:30, 13:30 to 14:00, 14:30  
3947 to 16:00;

---

3942 Find a time that works for everyone's schedule and constraints.  
3943 SOLUTION: Here is the proposed time: Monday, 9:00 - 9:30  
3944

3945 TASK: You need to schedule a meeting for Mason, Linda, Barbara, Roger and Lisa for  
3946 half an hour between the work hours of 9:00 to 17:00 on Monday.

3947 Here are the existing schedules for everyone during the day:  
3948 Masonhas no meetings the whole day.  
3949 Linda is busy on Monday during 9:30 to 11:00, 12:00 to 12:30, 13:00 to 13:30, 16:30  
3950 to 17:00;  
3951 Barbara has blocked their calendar on Monday during 9:00 to 9:30, 12:00 to 13:00,  
3952 15:30 to 16:00;  
3953 Roger has meetings on Monday during 9:00 to 9:30, 13:00 to 14:30, 15:00 to 17:00;  
3954 Lisa has blocked their calendar on Monday during 9:00 to 10:00, 10:30 to 11:30,  
3955 12:00 to 12:30, 13:00 to 14:30, 15:00 to 17:00;

3956 Mason would rather not meet on Monday after 13:30. Find a time that works for  
3957 everyone's schedule and constraints.  
3958 SOLUTION: Here is the proposed time: Monday, 11:30 - 12:00

3959 TASK: You need to schedule a meeting for Thomas, Olivia, Doris, Amber and Ronald for  
3960 half an hour between the work hours of 9:00 to 17:00 on Monday.

3961 Here are the existing schedules for everyone during the day:  
3962 Thomas has meetings on Monday during 12:00 to 12:30, 16:30 to 17:00;  
3963 Olivia has blocked their calendar on Monday during 10:30 to 11:30, 13:00 to 14:00,  
3964 16:00 to 17:00;  
3965 Doris is busy on Monday during 13:30 to 14:00, 16:30 to 17:00;  
3966 Amber is busy on Monday during 10:00 to 10:30, 12:00 to 13:00, 14:00 to 15:00, 15:30  
3967 to 17:00;  
3968 Ronald has meetings on Monday during 9:30 to 11:30, 12:00 to 13:00, 13:30 to 14:00,  
3969 15:00 to 15:30;

3970 Ronald would rather not meet on Monday before 11:30. Find a time that works for  
3971 everyone's schedule and constraints.  
3972 SOLUTION: Here is the proposed time: Monday, 11:30 - 12:00

3973 TASK: You need to schedule a meeting for Patrick, Kathleen, Frank, Patricia and  
3974 Carolyn for half an hour between the work hours of 9:00 to 17:00 on Monday.  
3975

3976 Here are the existing schedules for everyone during the day:  
3977 Patrick is free the entire day.  
3978 Kathleen has meetings on Monday during 9:30 to 10:00, 15:00 to 15:30;  
3979 Frank has blocked their calendar on Monday during 11:00 to 11:30, 15:30 to 17:00;  
3980 Patricia is busy on Monday during 9:00 to 12:00, 14:30 to 15:00, 16:00 to 16:30;  
3981 Carolyn has meetings on Monday during 9:00 to 9:30, 11:00 to 12:30, 13:00 to 17:00;

3982 Find a time that works for everyone's schedule and constraints.  
3983 SOLUTION: Here is the proposed time: Monday, 12:30 - 13:00

3984 Query:

3985 TASK: You need to schedule a meeting for Stephen, Elijah, William, Jeremy and  
3986 Timothy for half an hour between the work hours of 9:00 to 17:00 on Monday.

3987 Here are the existing schedules for everyone during the day:  
3988 Stephen is free the entire day.  
3989 Elijah is busy on Monday during 9:00 to 9:30, 12:30 to 13:00, 14:30 to 15:00, 16:00  
3990 to 16:30;  
3991 William is busy on Monday during 9:30 to 10:00, 15:30 to 16:00;  
3992 Jeremy is busy on Monday during 9:00 to 9:30, 10:00 to 12:00, 13:00 to 15:00, 15:30  
3993 to 17:00;  
3994 Timothy is busy on Monday during 10:00 to 10:30, 11:30 to 14:30, 15:30 to 16:00;

3995 Find a time that works for everyone's schedule and constraints.  
SOLUTION:

---

3996  
3997  
3998  
3999  
4000  
4001  
4002  
4003  
4004  
4005  
4006  
4007  
4008  
4009  
4010  
4011  
4012  
4013  
4014  
4015  
4016  
4017  
4018  
4019  
4020  
4021  
4022  
4023  
4024  
4025  
4026  
4027  
4028  
4029  
4030  
4031  
4032  
4033  
4034  
4035  
4036  
4037  
4038  
4039  
4040  
4041  
4042  
4043  
4044  
4045  
4046  
4047  
4048  
4049

=====RESPONSE OF o1-mini=====  
SOLUTION: Here is the proposed time: Monday, 15:00 - 15:30