

Example of VRP Prompting

- **Task Description:** You are given a list of customers with coordinates:

– (1): (87, 10)

– (2): (10, 87)

– (3): (87, 87)

– (4): (10, 10)

and a depot at (50, 50).

– (1): (87, 10)

– (2): (10, 87)

– (3): (87, 87)

– (4): (10, 10)

There are 50 customers in total.

of 50. The truck has a capacity of 50. It must visit

all the customers, starting and ending at the depot.

Can we use LLMs
for CPs?

and
nation

with capacity
to go through



Example of VRP Prompting

- **Task Description:** You are given a list of customers with coordinates:

- (1): (87, 39);
- (2): (1, 90);
- (3): (28, 65);
- (4): (85, 85);

Positions



and a list of customer demands:

- (1): 12;
- (2): 8;
- (3): 16;
- (4): 5.

Demands



Objective and
Extra Information



There is a depot with coordinates (50, 50) and one vehicle with a maximum capacity of 50. The goal is to find the route that has the minimum total length and go through all the customers, starting and ending at the depot.

Example of VRP Prompting

- **Task Description:** You are given a list of customers with coordinates:

– (1): (87, 12)

– (2): (12, 87)

– (3): (12, 12)

– (4): (87, 87)

and a depot at (50, 50).

– (1):

– (2):

– (3):

– (4):

There are 100 customers

of 50. The truck

can carry 100 units and must visit

all the customers, starting and ending at the depot.

Chain-of-Thought?
Few-Shot?

and
ation

capacity
go through



Example of VRP Prompting

- **Task Description:** You are given a list of customers with coordinates:

– (1): (87, 12)

– (2): (45, 67)

– (3): (23, 89)

– (4): (56, 34)

and a depot at (0, 0).

– (1): (87, 12)

– (2): (45, 67)

– (3): (23, 89)

– (4): (56, 34)

There is a vehicle with a capacity

of 50. The vehicle must visit

all the customers, starting and ending at the depot.

Can we come up with
better strategy?

and
ation

capacity
go through



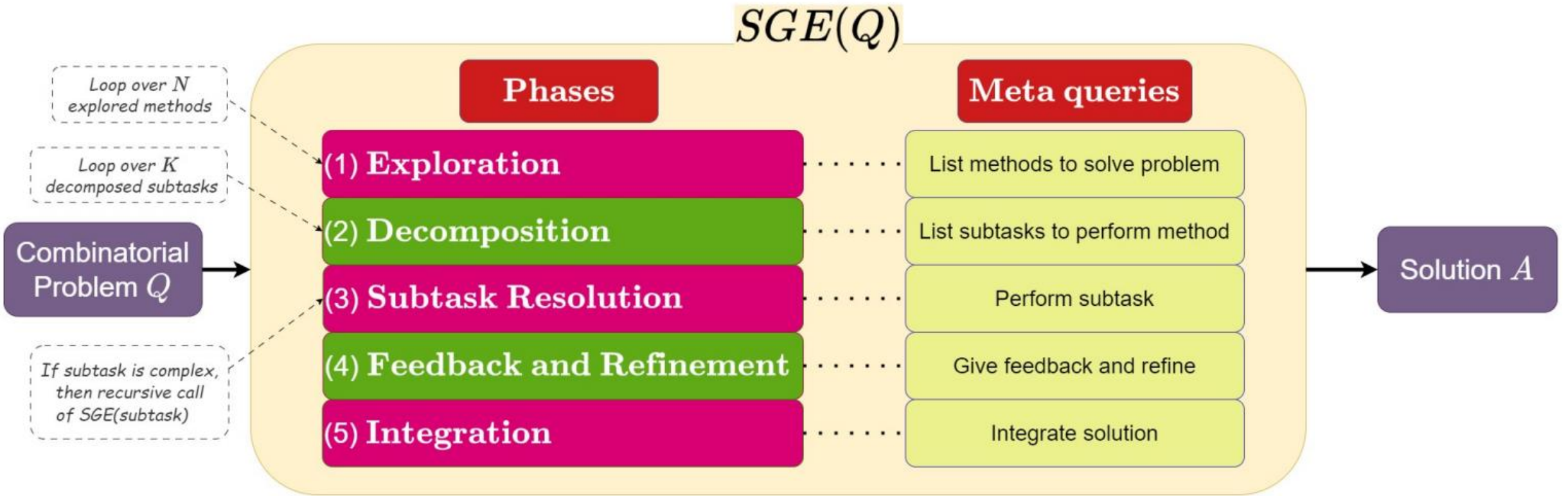


MOHAMED BIN ZAYED
UNIVERSITY OF
ARTIFICIAL INTELLIGENCE

Self-Guiding Exploration for Combinatorial Problems (NeurIPS24)

Zangir Iklassov, Yali Du, Farkhad Akimov, and Martin Takac

SGE - Overview

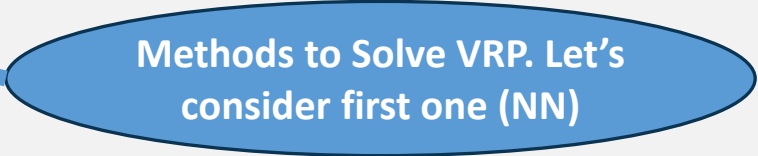


SGE - Exploration

Example of VRP Exploration

- **Prompt:** [Task Description]. List heuristic methods to solve this problem. Return only method names separated by new lines.
- **Exploration Output:**
 - Nearest Neighbor Algorithm
 - Savings Algorithm
 - Sweep Algorithm
 - Clarke and Wright Algorithm
 - Genetic Algorithm
 - Simulated Annealing
 - Tabu Search
 - Ant Colony Optimization
 - Particle Swarm Optimization

Methods to Solve VRP. Let's consider first one (NN)



SGE – Decomposition

Steps of NN method

Example of VRP Decomposition

- **Prompt:** [Task Description]. Nearest Neighbor Algorithm. List the steps to use this heuristic method. Return only the steps, separated by new lines.
- **Decomposition Output:**
 - ****Initialize at the Depot:**** The starting point is the depot located at coordinates (50, 50).
 - ****Find the Nearest Customer:**** Calculate the Euclidean distance from the depot to each customer, and move to the closest one.
 - ****Mark the Customer as Visited:**** Once a customer is visited, they should not be revisited.
 - ****Repeat Finding and Moving:**** Continue finding the nearest unvisited customer from the current location until all customers have been visited.
 - ****Return to the Depot:**** After visiting the last customer, calculate the distance back to the depot to complete the route.
 - ****Calculate the Total Length of the Route:**** Sum up all the distances traveled from start to finish.

SGE – Subtask Resolution

Example of Subtask Resolution, Step One

- **Prompt:** [Task Description]. [Apply the heuristic's step one].
- **Subtask Resolution Output:**

Initialize at the Depot

- Current location: Depot at (50, 50).

NN step 1 output.



SGE – Subtask Resolution

Example of Subtask Resolution, Step Two

- **Prompt:** [Task Description]. [Apply the heuristic's step two].
- **Subtask Resolution Output:**

Find the Nearest Customer

- Distance to Customer 1 at (87, 39): $\sqrt{(87 - 50)^2 + (39 - 50)^2}$
- Distance to Customer 2 at (1, 90): $\sqrt{(1 - 50)^2 + (90 - 50)^2}$
- Distance to Customer 3 at (28, 65): $\sqrt{(28 - 50)^2 + (65 - 50)^2}$
- Distance to Customer 4 at (85, 85): $\sqrt{(85 - 50)^2 + (85 - 50)^2}$

The initial distances from the depot to each customer are as follows:

- Customer 1: 38.60
- Customer 2: 63.25
- Customer 3: 26.63 (nearest)
- Customer 4: 49.50

NN step 2 output

SGE – Subtask Resolution

Example of Subtask Resolution, Step Three

- **Prompt:** [Task Description]. [Apply the heuristic's step three].
- **Subtask Resolution Output:**

Mark the Customer as Visited.

The nearest customer is Customer 3 at coordinates (28, 65). We will visit Customer 3 first.

NN step 3 output.
Steps 4 and 5 are similar
just done in loop

SGE – Subtask Resolution

Example of Subtask Resolution, Step Six

- **Prompt:** [Task Description]. [Apply the heuristic's step six].
- **Subtask Resolution Output:**

Calculate the Total Length of the Route.

The complete route, following the Nearest Neighbor Algorithm from the depot, visiting all customers, and returning to the depot, totals a distance of approximately 232.22 units.

Here's a summary of the route taken:

- Start at Depot (50, 50)
- Visit Customer 3 (28, 65)
- Visit Customer 2 (1, 90)
- Visit Customer 4 (85, 85)
- Visit Customer 1 (87, 39)
- Return to Depot (50, 50)



NN step 6 output
Candidate solution

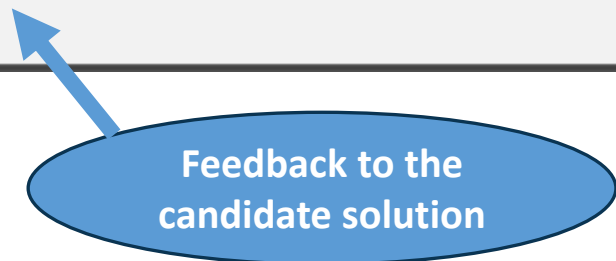
SGE – Feedback

Example of Feedback Prompting

- **Prompt:** [Task Description]. [Nearest Neighbor candidate solution]. Give feedback to the proposed solution.

- **Feedback Output:**

Implement route optimization post-processing, such as 2-opt or 3-opt techniques, which iteratively refine the route by removing crossings and reducing travel distance.

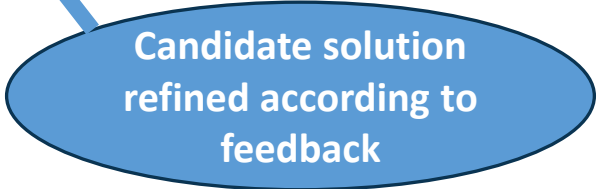


Feedback to the candidate solution

SGE – Refinement

Example of Refinement Prompting

- **Prompt:** [Task Description]. [Nearest Neighbor candidate solution]. [Feedback]. Improve the solution given feedback.
- **Refinement Output:**
 - we'll apply a 2-opt optimization to the route. This will involve swapping segments of the route to find a shorter path. Given the route '1 -> 2 -> 3 -> 4', we'll try different swaps and see if we can reduce the total distance further.
 - Let's simulate the 2-opt optimization on this route.
 - After applying the 2-opt optimization technique, the final optimized route has been adjusted to: Customer 1 -> Customer 3 -> Customer 2 -> Customer 4, then returning to the depot. This rearrangement was made to minimize the total distance traveled.



Candidate solution refined according to feedback

SGE – Integration

Example of Solution Integration

- **Prompt:** [Task Description]. [All candidate solutions]. Integrate all previous findings and provide the final answer. Return the answer only as a Python list of elements.
- **Refinement Output:** [1, 3, 2, 4]



Integration of all
candidate solutions

Setup

Tasks
Assignment Problem
Knapsack Problem
Bin Packing Problem
Traveling Salesman Problem
Vehicle Routing Problem
Job Scheduling Problem

Baselines
Input-Output (IO) Direct Prompting
Chain-of-Thought Prompting
Self-Refine (Refine) Prompting
Decomposition Prompting

$$100 \times \frac{g_{io} - g_{method}}{g_{io}}$$

Metrics: Improvement of cost function - g with respect to IO solution

$$100 \times \frac{g_{method} - g_{opt}}{g_{opt}}$$

Metrics: gap with optimal solution

Results on CP

Task	GPT-4				Gemini-1.5			
	CoT	Refine	Decomp	Ours	CoT	Refine	Decomp	Ours
Assignment	11.46	14.47	33.80	41.33	11.66	13.98	31.94	40.46
Knapsack	15.37	17.16	51.95	70.39	13.85	16.85	48.62	65.87
Bin Packing	14.06	17.12	39.57	74.72	11.89	15.43	35.74	67.63
Travelling Salesman	13.64	15.75	38.49	72.10	14.34	15.90	36.36	68.09
Vehicle Routing	14.27	16.94	36.73	71.92	11.88	15.13	33.59	68.02
Job Scheduling	13.84	16.37	38.20	75.33	13.41	15.75	36.36	67.89

Optimality gap.
The lower it is , the better

Improvement of cost w.r.t IO.
The higher it is , the better

Size	Method	Assignment	Knapsack	Bin Packing	TSP	VRP	JSP
5 NODES	IO	45.45	90.10	108.2	100.3	102.0	105.3
	CoT	39.33	66.88	78.24	81.15	78.17	79.41
	Refine	36.42	61.98	77.40	71.62	72.49	71.72
	Decomp	14.66	21.56	40.00	43.62	40.65	44.15
	Ours	2.500	8.050	9.060	8.27	11.92	9.300
8 NODES	IO	46.84	103.5	112.8	116.9	116.3	108.2
	CoT	39.70	73.84	85.08	89.01	89.48	85.21
	Refine	37.32	72.62	86.25	85.59	83.31	78.43
	Decomp	18.49	26.43	52.73	53.48	54.43	49.81
	Ours	8.290	14.88	20.95	15.19	19.65	21.26
12 NODES	IO	49.11	101.5	120.7	121.6	118.5	117.6
	CoT	41.70	79.33	93.84	86.84	90.05	89.29
	Refine	40.35	77.09	82.23	88.57	88.40	87.02
	Decomp	21.12	35.82	55.40	57.51	59.19	56.01
	Ours	11.26	16.82	22.38	16.12	24.00	22.86

Results on Reasoning Tasks

Arithmetic tasks - GSM8K dataset

Problem: Beth bakes 4, 2 dozen batches of cookies in a week. If these cookies are shared amongst 16 people equally, how many cookies does each person consume?

Solution: Beth bakes 4 2 dozen batches of cookies for a total of $4*2 = \llcorner 4*2=8 \gg$ 8 dozen cookies

There are 12 cookies in a dozen

She splits the 96 cookies

Final Answer: 6

Problem: Mrs. Lim

she got 18 gallons

much was her re

Mrs. Lim got 68 g

So she was able

She was able to

Thus, her total re

Final Answer: 6

Problem: Tina b

of the people hav

Solution: Tina b

6 people attend t

Each of those pec

Two people drink 4

With one person drinking 3, that brings the total drank to $3+3+3+3 = \llcorner 3+3+3+3=25 \gg$ 25 sodas

As Tina started off with 36 sodas, that means there are $36-25 = \llcorner 36-25=11 \gg$ 11 sodas left

Final Answer: 11

Is it general-purpose method?



Results on Reasoning Tasks

Arithmetic tasks - GSM8K dataset

Problem: Beth bakes 4, 2 dozen batches of cookies in a week. If these cookies are shared amongst 16 people equally, how many cookies does each person consume?

Solution: Beth bakes 4 2 dozen batches of cookies for a total of $4*2 = \ll4*2=8\gg$ 8 dozen cookies
There are 12 cookies in a dozen and she makes 8 dozen cookies for a total of $12*8 = \ll12*8=96\gg$ 96 cookies
She splits the 96 cookies equally amongst 16 people so they each eat $96/16 = \ll96/16=6\gg$ 6 cookies

Final Answer: 6

Problem: Mrs. Lim milks her cows twice a day. Yesterday morning, she got 68 gallons of milk and in the evening, she got 82 gallons. This morning, she got 18 gallons fewer than she had yesterday morning. After selling some gallons of milk in the afternoon, Mrs. Lim has only 24 gallons left. How much was her revenue for the milk if each gallon costs \$3.50?

Mrs. Lim got 68 gallons - 18 gallons = $\ll68-18=50\gg$ 50 gallons this morning.
So she was able to get a total of 68 gallons + 82 gallons + 50 gallons = $\ll68+82+50=200\gg$ 200 gallons.
She was able to sell 200 gallons - 24 gallons = $\ll200-24=176\gg$ 176 gallons.
Thus, her total revenue for the milk is $\$3.50/\text{gallon} \times 176 \text{ gallons} = \ll3.50*176=616\gg$ 616.

Final Answer: 616

Problem: Tina buys 3 12-packs of soda for a party. Including Tina, 6 people are at the party. Half of the people at the party have 3 sodas each, 2 of the people have 4, and 1 person has 5. How many sodas are left over when the party is over?

Solution: Tina buys 3 12-packs of soda, for $3*12 = \ll3*12=36\gg$ 36 sodas
6 people attend the party, so half of them is $6/2 = \ll6/2=3\gg$ 3 people
Each of those people drinks 3 sodas, so they drink $3*3 = \ll3*3=9\gg$ 9 sodas
Two people drink 4 sodas, which means they drink $2*4 = \ll4*2=8\gg$ 8 sodas
With one person drinking 5, that brings the total drank to $5+9+8+3 = \ll5+9+8+3=25\gg$ 25 sodas
As Tina started off with 36 sodas, that means there are $36-25 = \ll36-25=11\gg$ 11 sodas left

Final Answer: 11

Results on Reasoning Tasks

Different Reasoning Tasks

Method	Arithmetic				Commonsense			Symbolic	Avg.
	AQUA	GSM8K	SVAMP	ASDiv	StrategyQA	CSQA	ARC	LastLetter	
IO Prompting	67.30	87.04	88.34	90.10	78.40	81.14	87.52	81.98	82.73
CoT Prompting	69.57	89.76	91.58	93.32	81.16	84.15	90.80	85.18	85.69
Refine Prompting	69.68	89.80	91.00	93.10	81.26	83.50	91.09	84.82	85.53
Decomp Prompting	69.99	91.85	92.16	94.08	82.08	84.88	91.76	85.64	86.43
Ours	74.63	97.35	98.16	97.24	83.49	85.68	93.28	86.96	89.60

Accuracy on test data. The higher it is , the better

Thank you!

