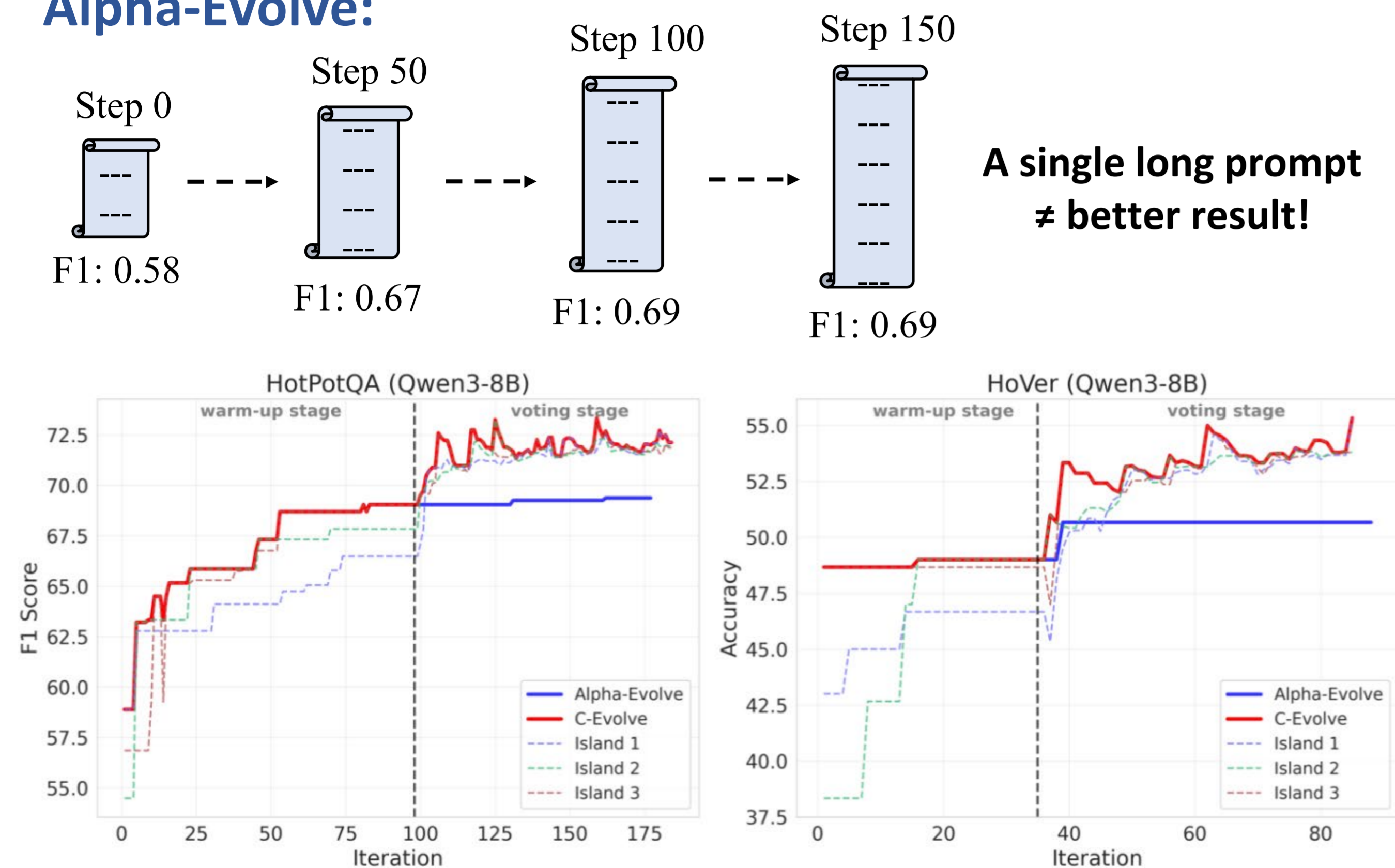
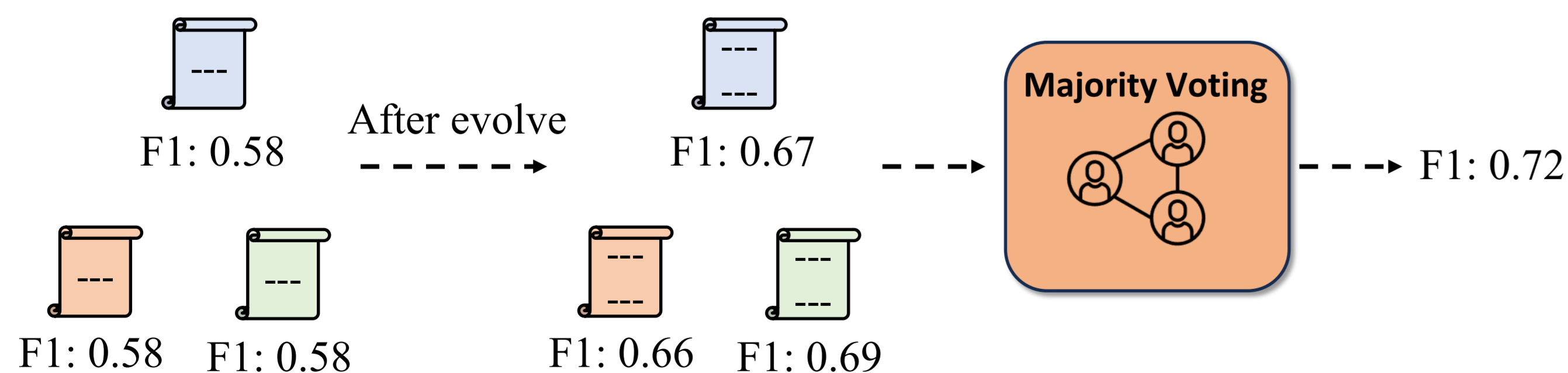


Alpha-Evolve:



Is it possible to move beyond the limitations of a single prompt and instead use Prompt Groups to achieve better performance?

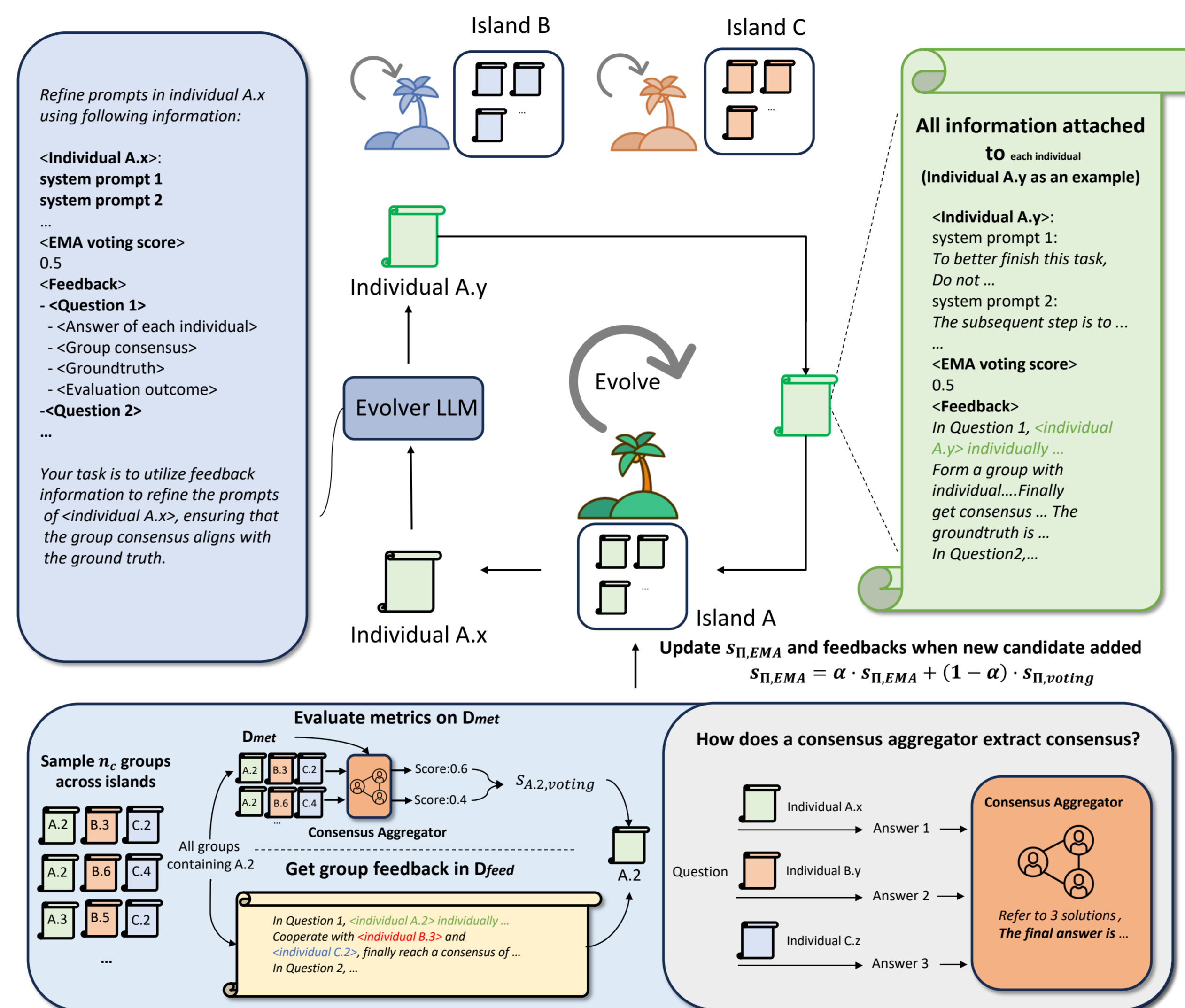
C-Evolve:



We did it! Our proposed method evolves a group of prompts that collectively achieve better performance after reaching consensus.

C-Evolve: Consensus-based Evolutionary Algorithm

Using a voting score to measure how well a prompt contributes when operating with others, rather than evaluating it in isolation



• Voting Score

Evaluate the contribution of each prompt within groups, which is calculated by averaging the performance of all the groups in which it participates

$$s_{\Pi,voting} = \frac{\sum_{k=1}^{n_c} \mathbb{I}(\Pi \in \mathcal{G}_k) \cdot \mathbb{E}_{(x,m) \sim \mathcal{D}_{met}} [u(C(y^{\mathcal{G}_k}), m)]}{\sum_{k=1}^{n_c} \mathbb{I}(\Pi \in \mathcal{G}_k)}$$

• EMA Voting Score

Encourage the selection and reproduction of individuals more likely to form a high-performing group, while eliminating under-performing ones.

$$s_{\Pi,EMA} \leftarrow \alpha \cdot s_{\Pi,EMA} + (1 - \alpha) \cdot s_{\Pi,voting}$$

Experimental results

- Achieves the state-of-the-art performance across a wide range of tasks

Model	HotpotQA	IFBench	Hover	MATH	GPQA	Improvement
Qwen3-8B						
Baseline	50.03	31.29	37.66	67.66	41.43	-
GEPA	65.72	34.01	43.33	-	-	+8.02
AlphaEvolve	65.31	41.15	44.66	82.66	43.08	+9.75
C-Evolve	70.67	43.88	50.33	85.33	47.15	+13.85
GPT-4.1 mini						
Baseline	44.24	40.13	42	78.66	46.34	-
GEPA	68.39	46.59	39	-	-	+9.2
AlphaEvolve	67.31	45.24	50.33	92.66	63.01	+13.42
C-Evolve	70.64	47.96	51.66	95.33	66.26	+16.09

- Outperforms the Single Prompt Evolution

Method	island 1	island 2	island 3	Final Accuracy
AlphaEvolve	41.15	36.73	37.07	41.15 (individual)
C-Evolve	40.47	39.79	38.77	41.15 (consensus)
				43.88

- Helps on Hard Problems

	level 1	level 2	level 3	level 4	level 5
island 1	95	94.67	90.66	82.66	66.66
island 2	96.67	96	90	80	63.66
island 3	95.66	95.66	90.33	82	62.66
C-Evolve	96.67	96	92	82.66	68.33

Mail to: {litiancheng, wangyuhang, chenzhiyang, wangzijun, maliyuan}@westlake.edu.cn
guojunq@gmail.com