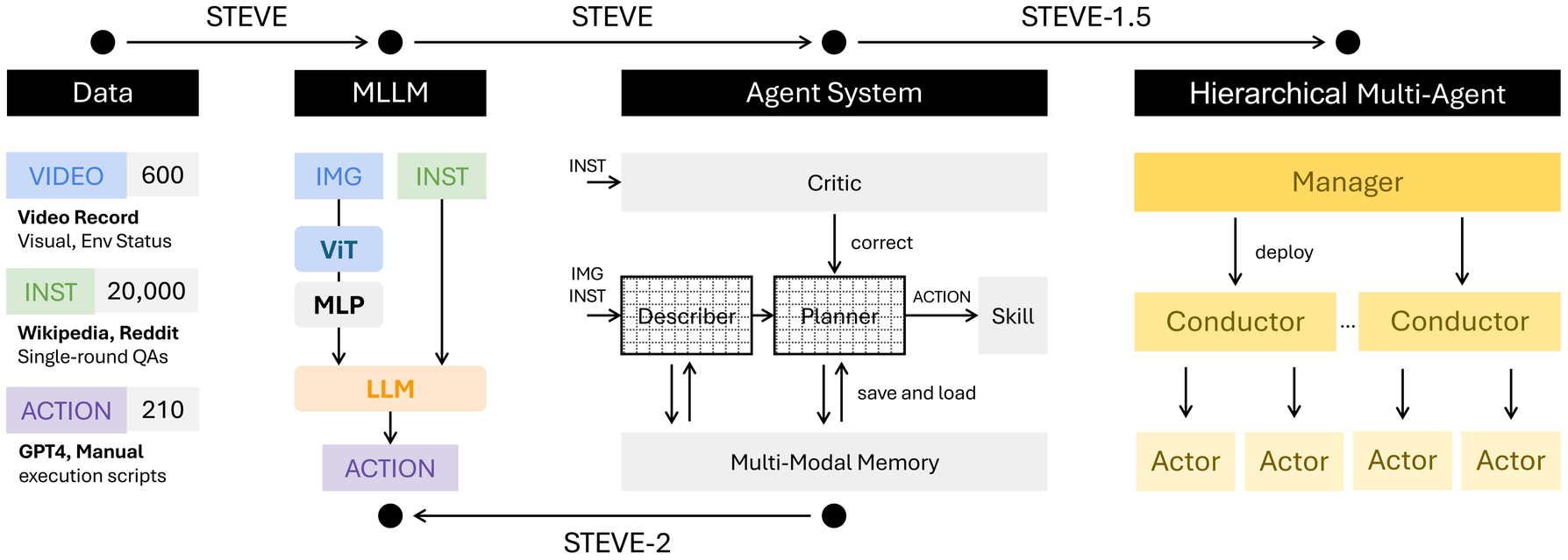
STEVE Series

Step-by-Step Construction of Agent Systems in Minecraft

Zhonghan Zhao[1*], Wenhao Chai[2*], Xuan Wang[1], Ke Ma[1], Kewei Chen[3], Dongxu Guo[3], Tian Ye[4], Yantin Zhang[3], Hongwei Wang[1], Gaoang Wang[1†]

1 Zhejiang University 2 University of Washington 3 Donghua University 4 Hong Kong University of Science and Technology (GZ)



ARCHITECTURE

TASK

STEVE	Basic Skill	STEVE-1.5	Navigation	STEVE-2	Creation	Future
Knowledge QA Tech tree mastery		Goal and block search Map exploration		Material collection Building creation		Take into the real world

Knowledge QA		Tech Tree Mastery	
Model	preference (↑)	Method	# iters (↓)
Llama2-13B [12]	6.89	AutoGPT [9]	107
GPT-4 [6]	8.04	Voyager [13]	35
STEVE-13B [17]	**8.12**	**STEVE-1** [17]	**33**

Table 1. **Comparison on Basic Skill**. Models preference rated 0-10 on knowledge QA and # iters stand for average iterations for

Method	# LLMs	Goal Search	Map Explore
		success (↑)	# area (↑)
Voyager [13]	12 / 20	64%	755
STEVE-1 [17]	20 / 24	64%	696
STEVE-2 [18]	5 / 8	**91%**	**1493**

Table 2. **Comparison on Navigation.** We list the success rate of Goal Search. # area is the average squares of blocks over 5 iterations. We list the best performance with the number of LLMs for different tasks.

Method	# LLMs	Material Collection	Building Creation
		completion (↑)	FID (↓)
Voyager [13]	4	72%	256.75
Creative Agents [15]	4	-	68.32
STEVE-2 [18]	8 / 2	**99%**	**21.12**

Table 3. **Comparison on Creation.** We list task completion rates and average FID scores for image quality. We list the best performance with the number of LLMs for different tasks.