

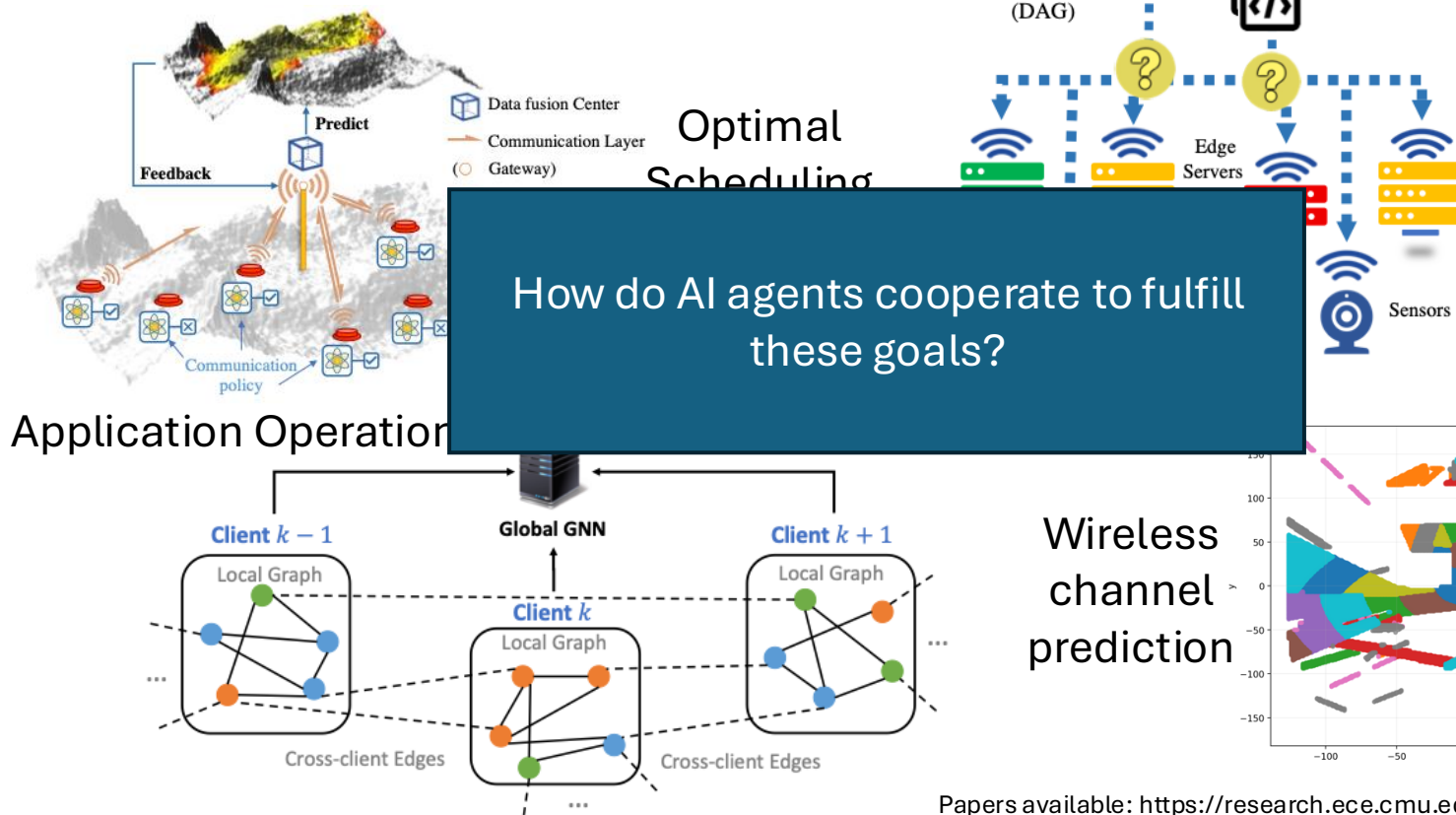
Quantifying Cooperation in Multi-Agent LLM Networks



Carlee Joe-Wong

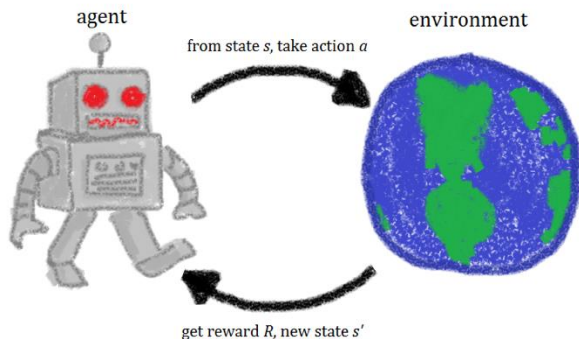
+ many others,
especially Hanqing Yang & Jingdi Chen

Learning in Networks

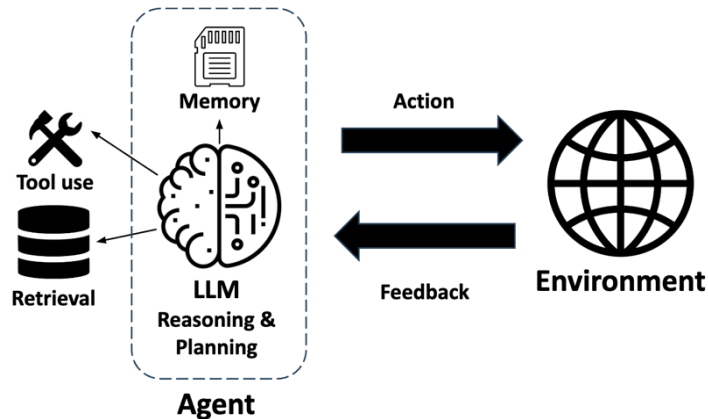


LLM “Agents” Are Powerful, But...

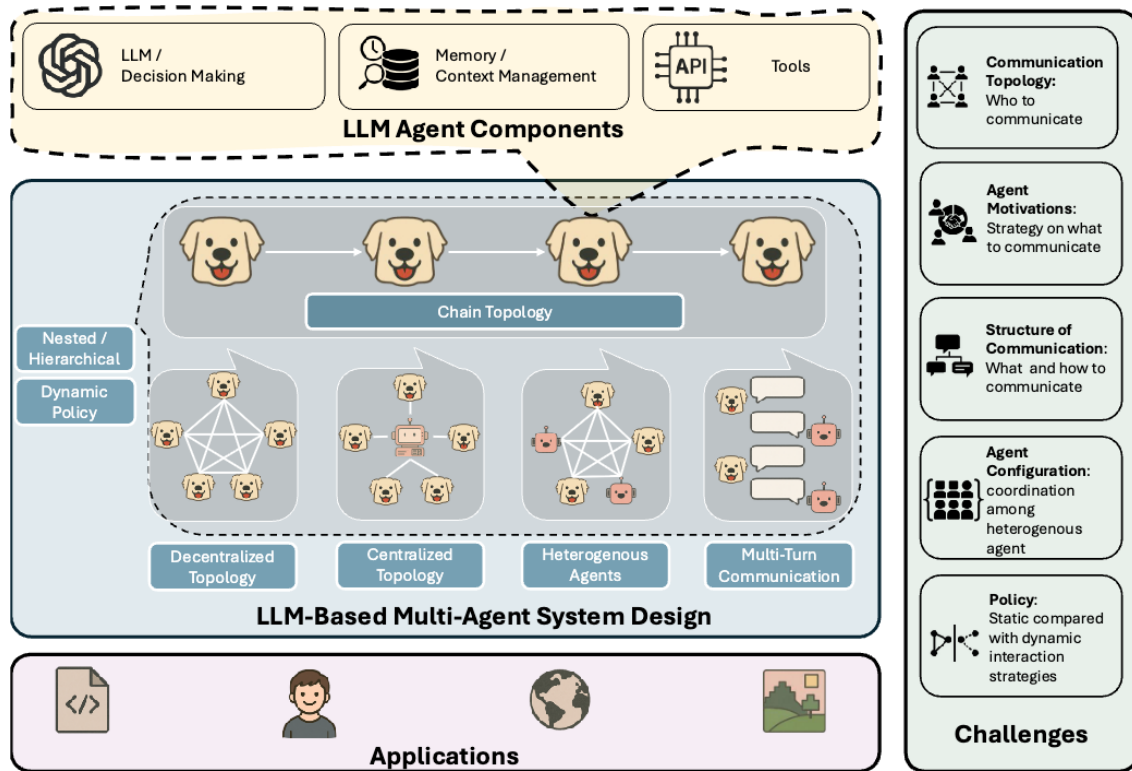
- “Agents” interact with some “environment”
 - Base stations in communication networks
 - Internet-of-Things devices



- LLM-based agents have:
 - Complex reasoning capabilities
 - Memory
 - ...but may be **unreliable**.



LLM Agent Collaboration?



CrewAI

<https://docs.crewai.com/en/enterprise/features/crew-studio>

n8n

<https://n8n.io/>

Langgraph

<https://docs.langchain.com/oss/python/langgraph/workflows-agents>

...and more!

What happens when you put a team of brilliant people together... and expect them to just work together?



Ignoring
dependency

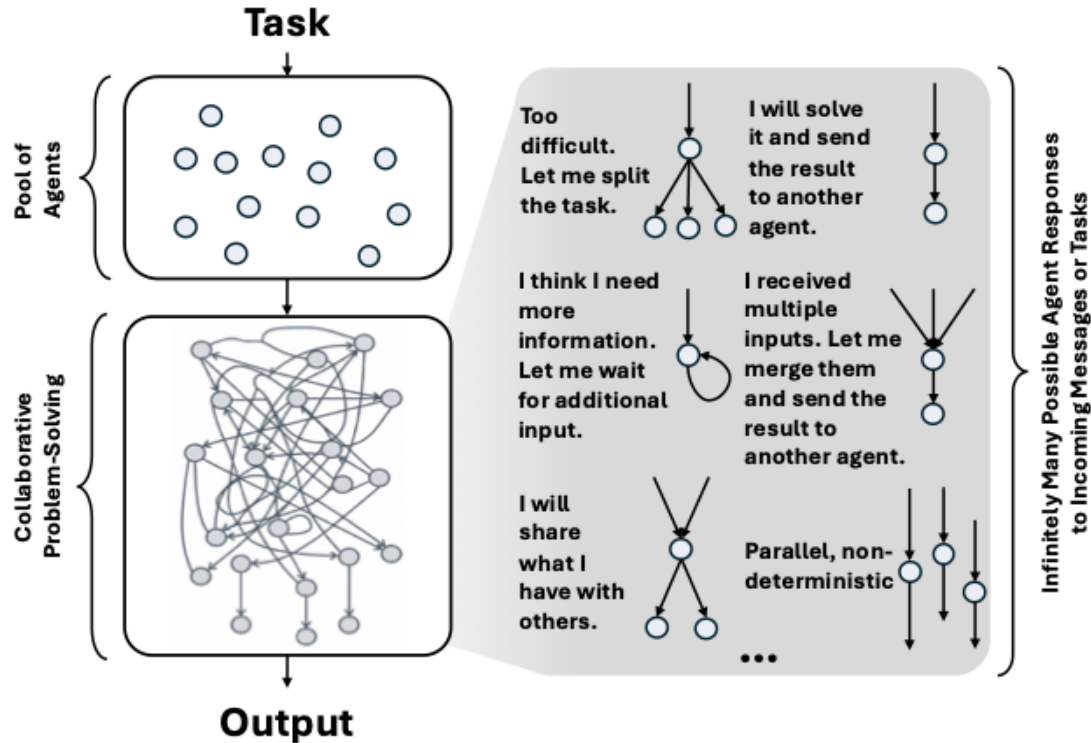
Overconfidence

Calculating
the same
thing twice

Not sure who to ask

Overloaded

What Collaboration Looks Like for Agents...



DIG to Heal

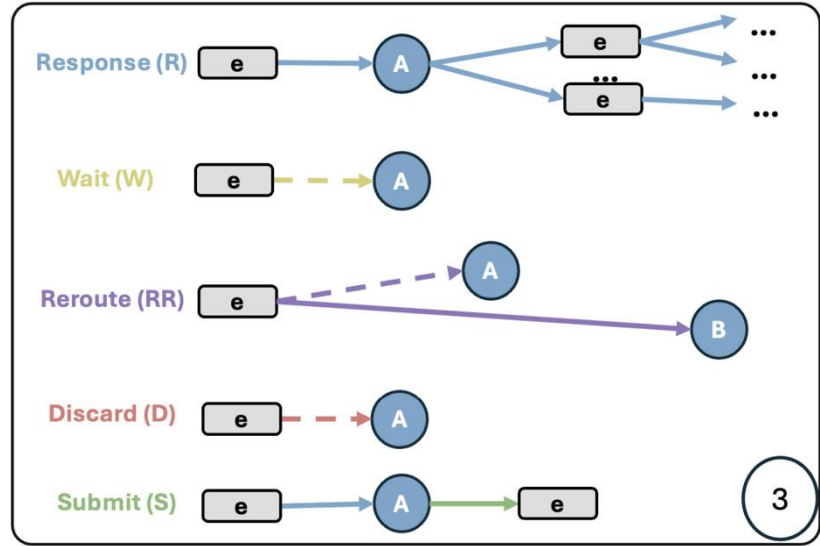
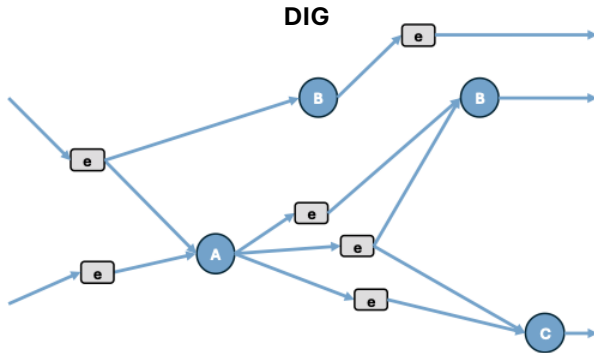
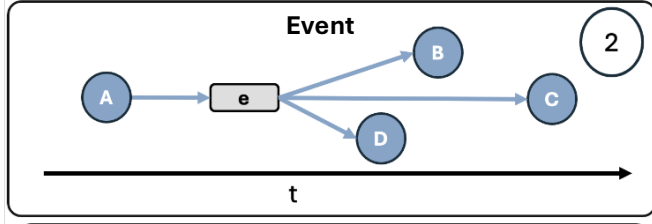
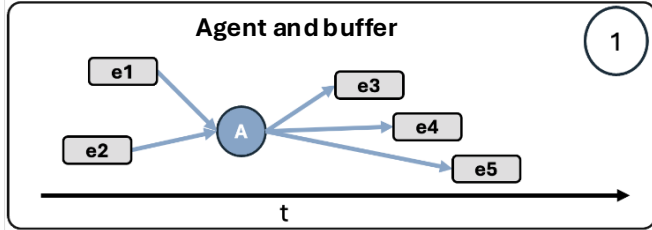
- We consider **cooperative problem-solving** in systems of multiple general-purpose agents
 - Asynchronous agent interactions
 - No predefined roles, control flow, or communication protocols
- **Dynamic Interactive Graph (DIG)** representation of agent interactions
 - Use the DIG structure to identify errors
 - Heal identified errors via agent feedback

Yang, H., Lee, H., Yao, Y., Liu, Z., Liu, K., Chen, J., & Joe-Wong, C. (2026). DIG to Heal: Scaling General-purpose Agent Collaboration via Explainable Dynamic Decision Paths. [arXiv preprint arXiv:2603.00309](https://arxiv.org/abs/2603.00309).

Code available: <https://happyeureka.github.io/dig/>

Dynamic Interaction Graph

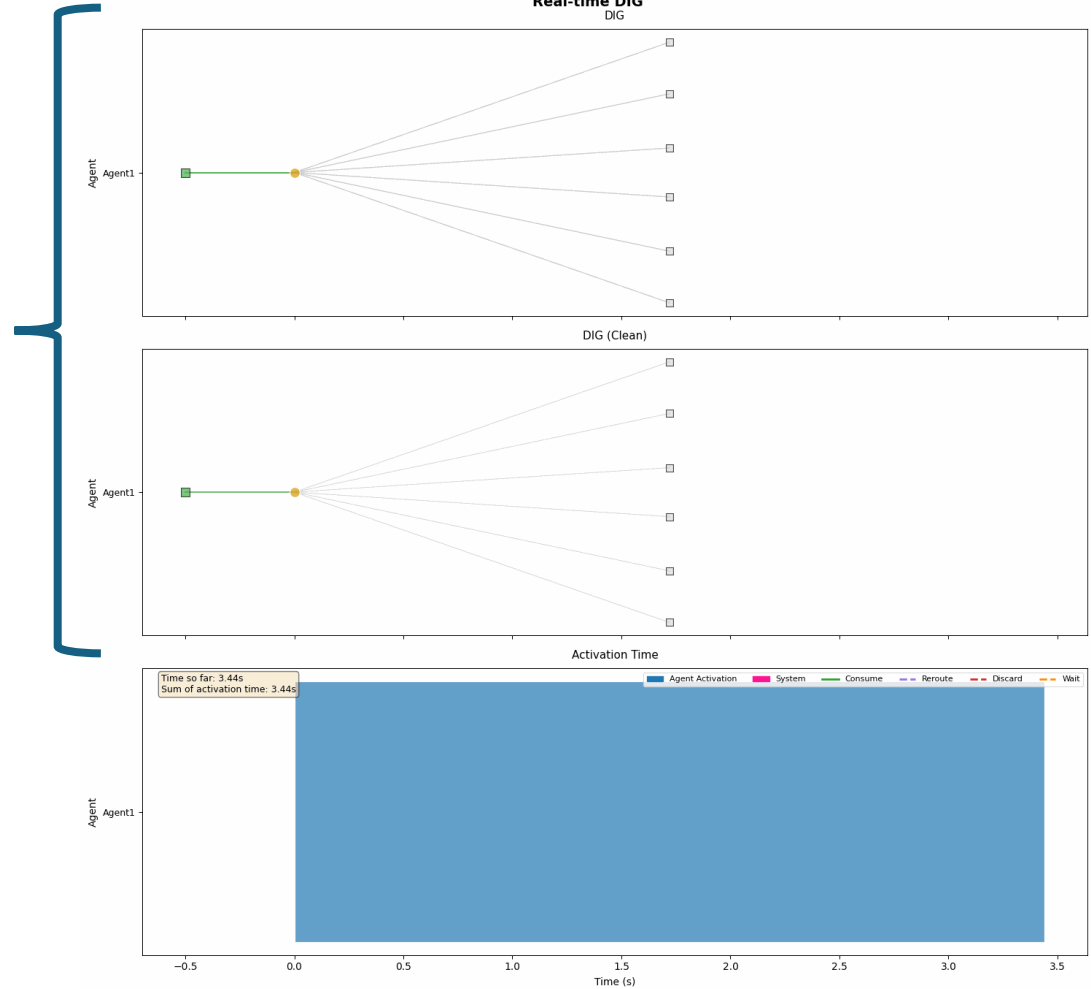
Agent responds to events by updating the DIG



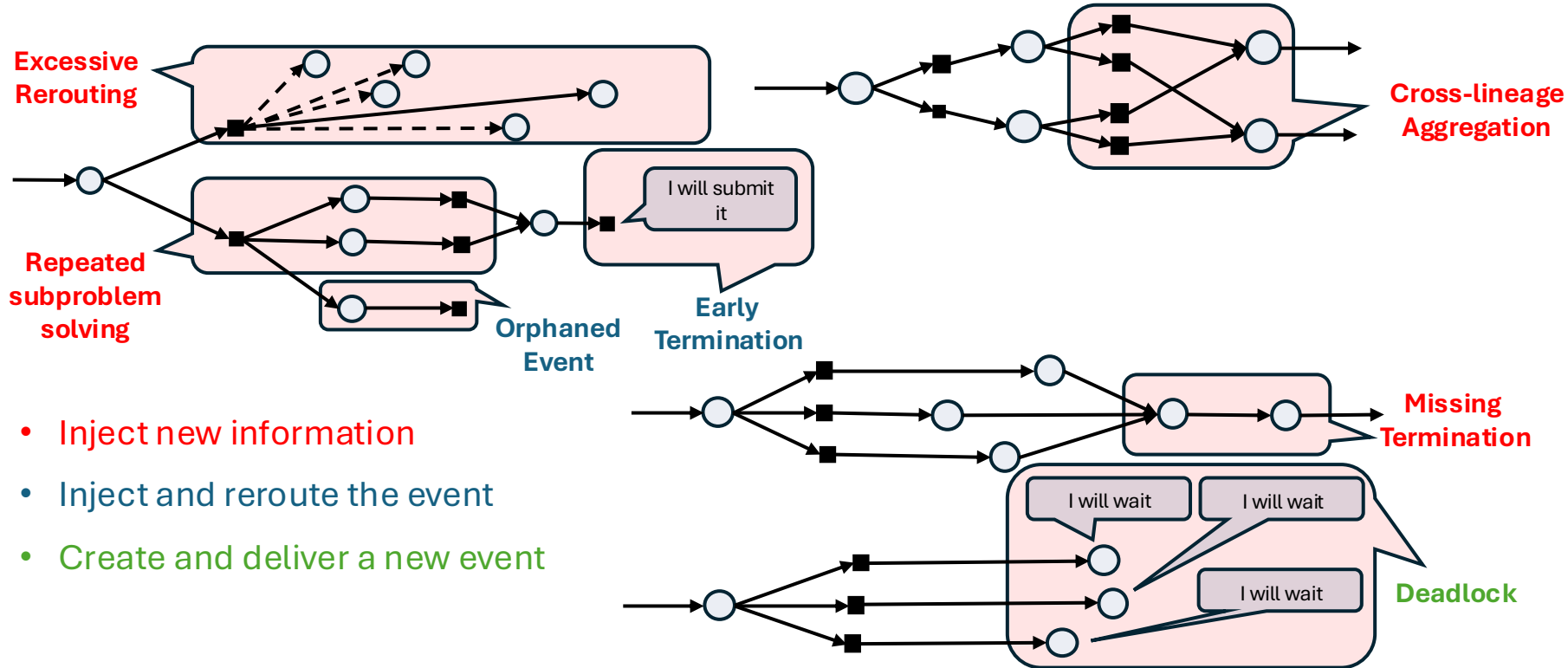
DIG in Action

DIG (with and without transient responses)

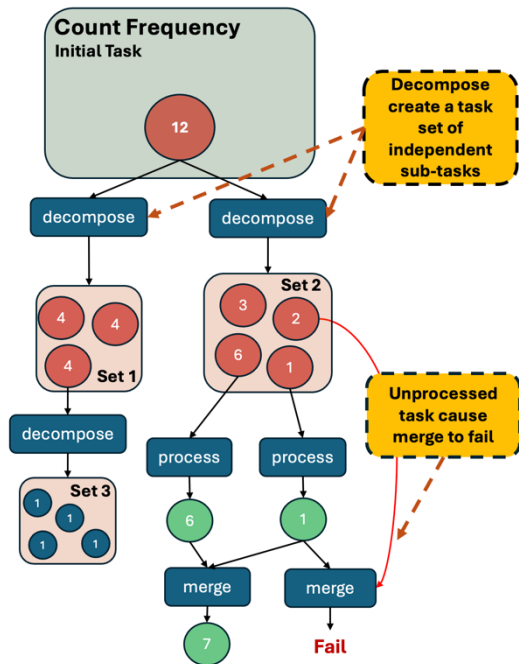
Agent activation
patterns with time



Detecting and Healing from Failures

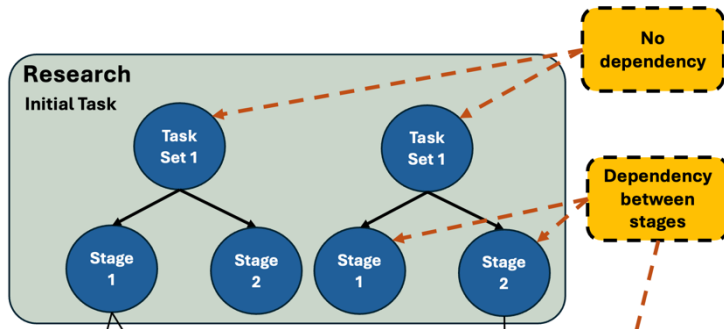


Healing Example



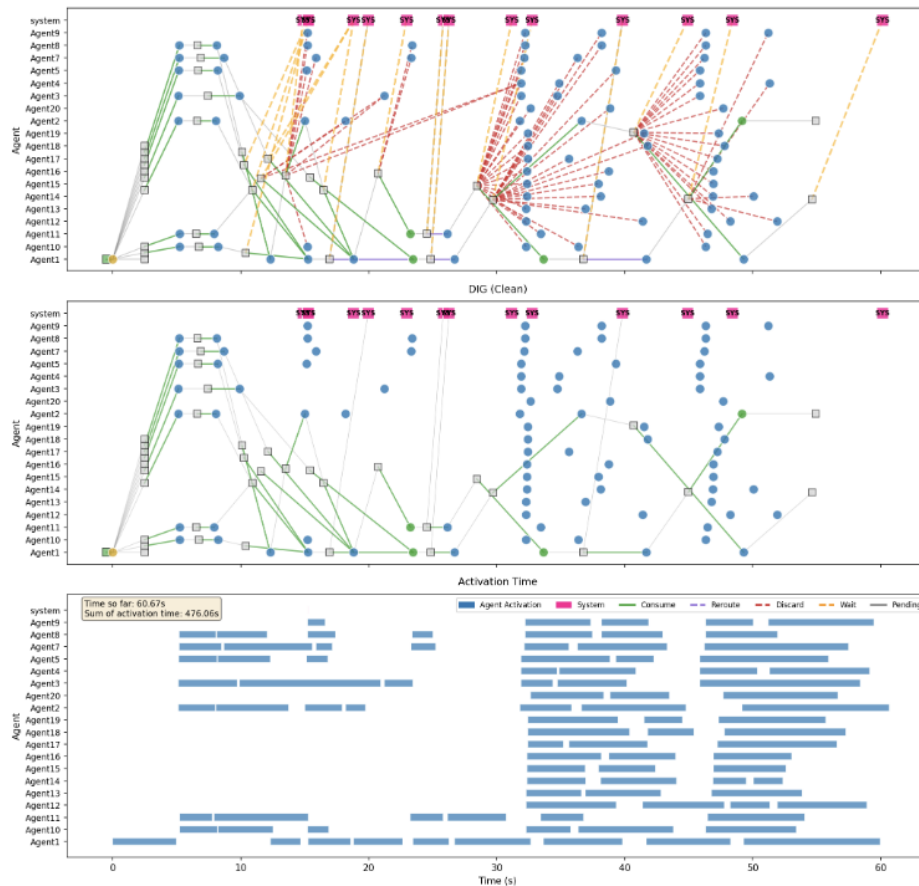
Setting	Agents	Problem Size	Detected Error ($\downarrow$)	Error ($\downarrow$)	Time ($\downarrow$)
Homogeneous (1000)					
MAS-Only	4	4000	2.11 ± 0.31	0.083 ± 0.118	13.0 ± 9.1
MAS+DIG	4	4000	2.60 ± 1.80	0.025 ± 0.075	10.7 ± 7.8
$\Delta\%$			+23%	-70%	-18%
MAS-Only	4	8000	2.60 ± 0.66	0.688 ± 0.084	14.3 ± 8.4
MAS+DIG	4	8000	9.10 ± 1.45	0.588 ± 0.138	20.1 ± 9.4
$\Delta\%$			+250%	-15%	+41%
MAS-Only	6	6000	2.50 ± 1.50	0.100 ± 0.249	9.2 ± 1.0
MAS+DIG	6	6000	2.60 ± 1.80	0.017 ± 0.050	10.4 ± 2.8
$\Delta\%$			+4%	-83%	+13%
MAS-Only	6	12000	3.50 ± 1.03	0.850 ± 0.033	19.4 ± 12.7
MAS+DIG	6	12000	8.80 ± 0.98	0.775 ± 0.065	19.7 ± 9.1
$\Delta\%$			+151%	-9%	+2%
Heterogeneous ([..., 1000,100,10,1])					
MAS-Only	4	1111	1.80 ± 0.40	0.000 ± 0.000	8.9 ± 0.9
MAS+DIG	4	1111	2.50 ± 1.86	0.001 ± 0.003	9.4 ± 2.5
$\Delta\%$			+39%	≈ 0	+6%
MAS-Only	4	2222	2.60 ± 0.80	0.735 ± 0.266	33.0 ± 7.9
MAS+DIG	4	2222	8.60 ± 1.63	0.533 ± 0.357	42.9 ± 8.6
$\Delta\%$			+231%	-27%	+30%
MAS-Only	6	111111	2.10 ± 0.30	0.100 ± 0.268	12.2 ± 8.1
MAS+DIG	6	111111	5.30 ± 3.49	0.092 ± 0.209	16.8 ± 9.4
$\Delta\%$			+152%	-8%	+38%
MAS-Only	6	222222	2.70 ± 1.10	0.847 ± 0.279	32.5 ± 13.2
MAS+DIG	6	222222	10.10 ± 1.64	0.568 ± 0.372	43.8 ± 11.0
$\Delta\%$			+274%	-33%	+35%

Scaling the DIG



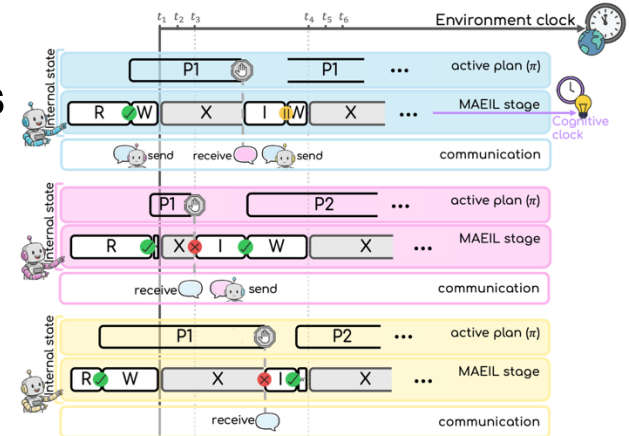
Setting	Detected Error ($\downarrow$)	Error ($\downarrow$)	Time ($\downarrow$)
Homogeneous (1000)			
MAS-Only	6.20 ± 2.96	0.50 ± 0.30	24.8 ± 31.8
MAS+DIG	8.40 ± 1.69	0.28 ± 0.28	17.9 ± 2.2
$\Delta\%$	+35%	-44%	-28%
Heterogeneous ([..., 1000,100,10,1])			
MAS-Only	3.50 ± 1.43	0.55 ± 0.15	43.4 ± 32.9
MAS+DIG	8.70 ± 1.85	0.45 ± 0.15	68.6 ± 26.7
$\Delta\%$	+149%	-18%	+58%

DIG with 20 agents



Future Work

- Modeling agent **planning and negotiation** as well as execution
- Optimizing agent **communication patterns**
- Monitoring for **task correctness**
- More general **failure detection & healing**
- More realistic **evaluation environments**



Thank you! Questions?

cjoewong@andrew.cmu.edu

<https://research.ece.cmu.edu/lions/index.html>