FinCon: A Synthesized LLM Multi-Agent System with Conceptual Verbal Reinforcement for Enhanced Financial Decision Making

Yangyang Yu*, Zhiyuan Yao*, Haohang Li*, Zhiyang Deng*, Yuechen Jiang*, Yupeng Cao*, Zhi Chen*, Jordan W. Suchow, Zhenyu Cui, Rong Liu, Zhaozhuo Xu, Denghui Zhang, Koduvayur Subbalakshmi, Guojun Xiong, Yueru He, Jimin Huang, Dong Li, Qianqian Xie

This is a joint work of
Stevens Institute of Technology, Harvard University, The FinAI

Introduction

Innovations of FinCon's design

LLMs remain insufficient for making high-quality financial sequential decisions. We propose FinCon, a comprehensive agent-based approach to using LLMs to manage financial decision making, including single-asset trading and portfolio management.

Gaps in literature	**Our solutions**
(1) Long-term risk exposure	A dual-level risk control component
(2) Inability for portfolio management	External portfolio optimization solver
(3) Heavy pressure on single-agent	A multi-agent system
(4) Peer-communication costs and unclear optimization goal	A synthesized Manager-Analyst hierarchical communication structure

Methodology

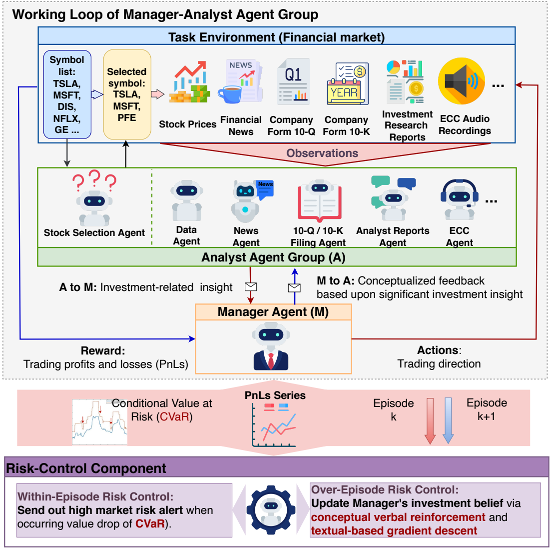
Methodology: trading tasks modeling

- We model two financial decision-making tasks as partially observable markov decision processes (POMDPs)
- The optimization objective for the whole system can be written as:

$$\max_{\boldsymbol{\theta}} \mathbb{E}\Big[\sum_{t=0}^{\infty} \alpha^t R_t^{\Pi^{\boldsymbol{\theta}}}\Big] \tag{1}$$

  where policies are parameterized by textual prompts $\boldsymbol{\theta}$.
- The FinCon system solves this optimization problem via a **Conceptual Verbal Reinforcement (CVRF) method** (a prompt engineering technique, without fine-tuning of language models' intrinsic parameters).

2.3 Synthesized multi-agent hierarchical structure design



Experiments

3.1 Experiments: single stock trading task

Comparison of key performance metrics during the testing period for the single-asset trading tasks involving **eight assets**, between FinCon and other algorithmic agents. [1]

Categories	Models	TSLA CR% ↑	TSLA SR↑	TSLA MDD% ↓	AMZN CR % ↑	AMZN SR↑	AMZN MDD% ↓	NIO CR% ↑	NIO SR↑	NIO MDD% ↓	MSFT CR% ↑	MSFT SR↑	MSFT MDD% ↓
Market	B&H	6.425	0.145	58.150	2.030	0.072	34.241	-77.210	-1.449	63.975	27.856	1.230	15.010
Our Model	FinCon	82.871	1.972	29.727	24.848	0.904	25.889	17.461	0.335	40.647	31.625	1.538	15.010
LLM-based	GA	16.535	0.391	54.131	-5.631	-0.199	37.213	-3.176	-1.574	3.155	-31.821	-1.414	39.808
	FinGPT	1.549	0.044	42.400	-29.811	-1.810	29.671	-4.959	-0.121	37.344	21.535	1.315	16.503
	FinMem	34.624	1.552	15.674	-18.011	-0.773	36.825	-48.437	-1.180	64.144	-22.036	-1.247	29.435
	FinAgent	11.960	0.271	55.734	-24.588	-1.493	33.074	0.933	0.051	19.181	-27.534	-1.247	39.544
DRL-based	A2C	-35.644	-0.805	61.502	-12.560	-0.444	37.106	-91.910	-1.728	68.911	21.397	0.962	21.458
	PPO	1.409	0.032	49.740	3.863	0.138	28.085	-72.119	-1.352	62.093	-4.761	-0.214	30.950
	DQN	-1.296	-0.029	58.150	11.171	0.398	31.174	-35.419	-0.662	56.905	27.021	1.216	21.458

Categories	Models	AAPL CR% ↑	AAPL SR↑	AAPL MDD% ↓	GOOG CR % ↑	GOOG SR↑	GOOG MDD% ↓	NFLX CR% ↑	NFLX SR↑	NFLX MDD% ↓	COIN CR% ↑	COIN SR↑	COIN MDD% ↓
Market	B&H	22.315	1.107	20.659	22.420	0.891	21.191	57.338	1.794	20.926	-21.756	-0.311	60.187
Our Model	FinCon	27.352	1.597	15.266	25.077	1.052	17.530	69.239	2.370	20.792	57.045	0.825	42.679
LLM-based	GA	5.694	0.372	14.161	-1.515	-0.192	8.210	41.770	1.485	20.926	19.271	0.277	67.532
	FinGPT	20.321	1.161	16.759	0.242	0.011	26.984	11.925	0.472	20.201	-99.553	-1.807	74.967
	FinMem	12.397	0.994	11.268	0.311	0.018	21.503	-10.306	-0.478	27.692	0.811	0.017	50.390
	FinAgent	20.757	1.041	19.896	-7.440	-1.024	10.360	61.303	1.960	20.926	-5.971	-0.106	56.882
DRL-based	A2C	13.781	0.683	14.226	8.562	0.340	21.191	-8.176	-0.258	49.579	-	-	-
	PPO	14.041	0.704	22.785	2.434	0.097	25.202	-33.144	-1.049	33.377	-	-	-
	DQN	21.125	1.048	16.131	20.690	0.822	21.191	21.753	0.687	39.733	-	-	-

[1] CR stands for Cumulative Return. SR stands for Sharpe Ratio. MDD stands for Max Drawdown. Note that the highest and second highest CRs and SRs have been tested and found statistically significant using the Wilcoxon signed-rank test. The highest CRs and SRs are highlighted in red, while the second highest are marked in blue.

3.2 Experiments: portfolio management task

FINCON's performance with the Markowitz Mean-Variance (MV) portfolio and FINRL in managing two sets of compact portfolios: Portfolio 1:[TSLA, MSFT, PFE]; Portfolio 2: [AMZN, GM, LLY]. FinCon leads the key performance metrics, CR and SR.

Models	**CR % ↑**	**SR↑**	**MDD %↓**
FinCon	113.836	3.269	16.163
Markowitz MV	12.636	0.614	17.842
FinRL-A2C	19.461	0.831	26.917
Equal-Weighted ETF	9.344	0.492	21.223

Table 1.1: Key performance metrics comparison among all portfolio management strategies of Portfolio 1.

Models	**CR % ↑**	**SR↑**	**MDD %↓**
FinCon	32.922	1.371	21.502
Markowitz MV	10.289	0.540	25.099
FinRL-A2C	11.589	0.649	15.787
Equal-Weighted ETF	15.061	0.867	14.662

Table 1.2: Key performance metrics comparison among all portfolio management strategies of Portfolio 2.

3.3 Experiments: ablation studies I

We evaluate our unique risk control component through two ablation studies. The first study assesses its ability to control risk within episodes using conditional value at risk (CVaR).

The performance of FinCon with the implementation of CVaR won a leading performance in both single-asset trading and portfolio management tasks.

Task	**Assets**	**Market Trend**	**Models**	**CR %↑**	**SR↑**	**MDD %↓**
Single Stock	GOOG	General Bullish ↗	w/ CVaR	25.077	1.052	17.530
			w/o CVaR	-1.461	-0.006	27.079
	NIO	General Bearish ↘	w/ CVaR	17.461	0.335	40.647
			w/o CVaR	-52.887	-1.002	70.243
Portfolio Management	(TSLA, MSFT, PFE)	Mixed	w/ CVaR	113.836	3.269	16.163
			w/o CVaR	14.699	1.142	17.511

Table 2: Key metrics FinCon with vs. without implementing **CVaR for within-episode risk control.**

3.3 Experiments: ablation studies II

The second study highlights the importance of the over-episode risk control mechanism in updating the trading manager agent's beliefs for a holistic understanding of current trading circumstances.

The performance of FinCon with the implementation of CVRF won a leading performance in both single-asset trading and portfolio management tasks.

Task	**Assets**	**Market Trend**	**Models**	**CR %↑**	**SR↑**	**MDD %↓**
Single Stock	GOOG	General Bullish ↗	w/ belief	25.077	1.052	17.530
			w/o belief	-11.944	-0.496	29.309
	NIO	General Bearish ↘	w/ belief	17.461	0.335	40.647
			w/o belief	8.197	0.156	55.688
Portfolio Management	(TSLA, MSFT, PFE)	Mixed	w/ belief	113.836	3.269	16.163
			w/o belief	28.432	1.181	27.535

Table 3: Key metrics FinCon with vs. without implementing **belief updates for over-episode risk control.**

4. Conclusion

- FinCon is a novel LLM-based multi-agent framework for financial decision-making tasks, including single stock trading and portfolio management.
- Central to the system is the Synthesized Manager-Analyst hierarchical communication structure and a dual-level risk control component.
- Manager agent synthesizes insights from specialized analyst agents who distill financial data from multiple sources into key investment insights.
- The dual-level risk control component introduces a new approach to defining agent personas, enabling dynamic updates of systematic risk and market beliefs within agent communication.

THANK YOU !