



Corso di Laurea in Informatica

Synaptic Coder

**Un sistema LLM multi-agente per la
generazione automatica di codice
sorgente**

Prof. Fabio Palomba
Dott. Antonio Della Porta

Dorotea Serrelli
Mat.: 0512113740



d.serrelli1@studenti.unisa.it



<https://github.com/DoroteaSerrelli>

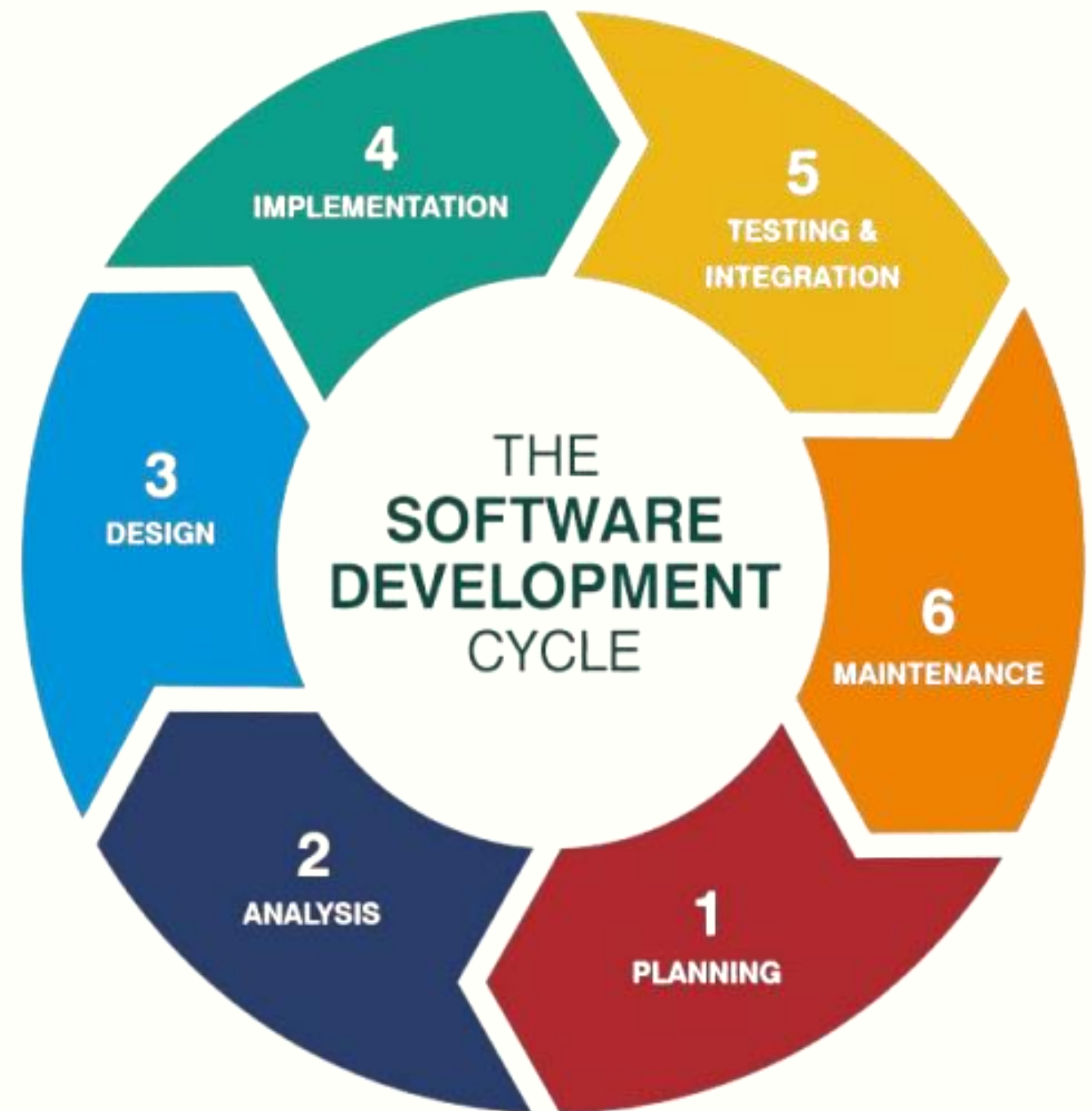


www.linkedin.com/in/dorotea-serrelli-a6b619370

PDF Tesi —>

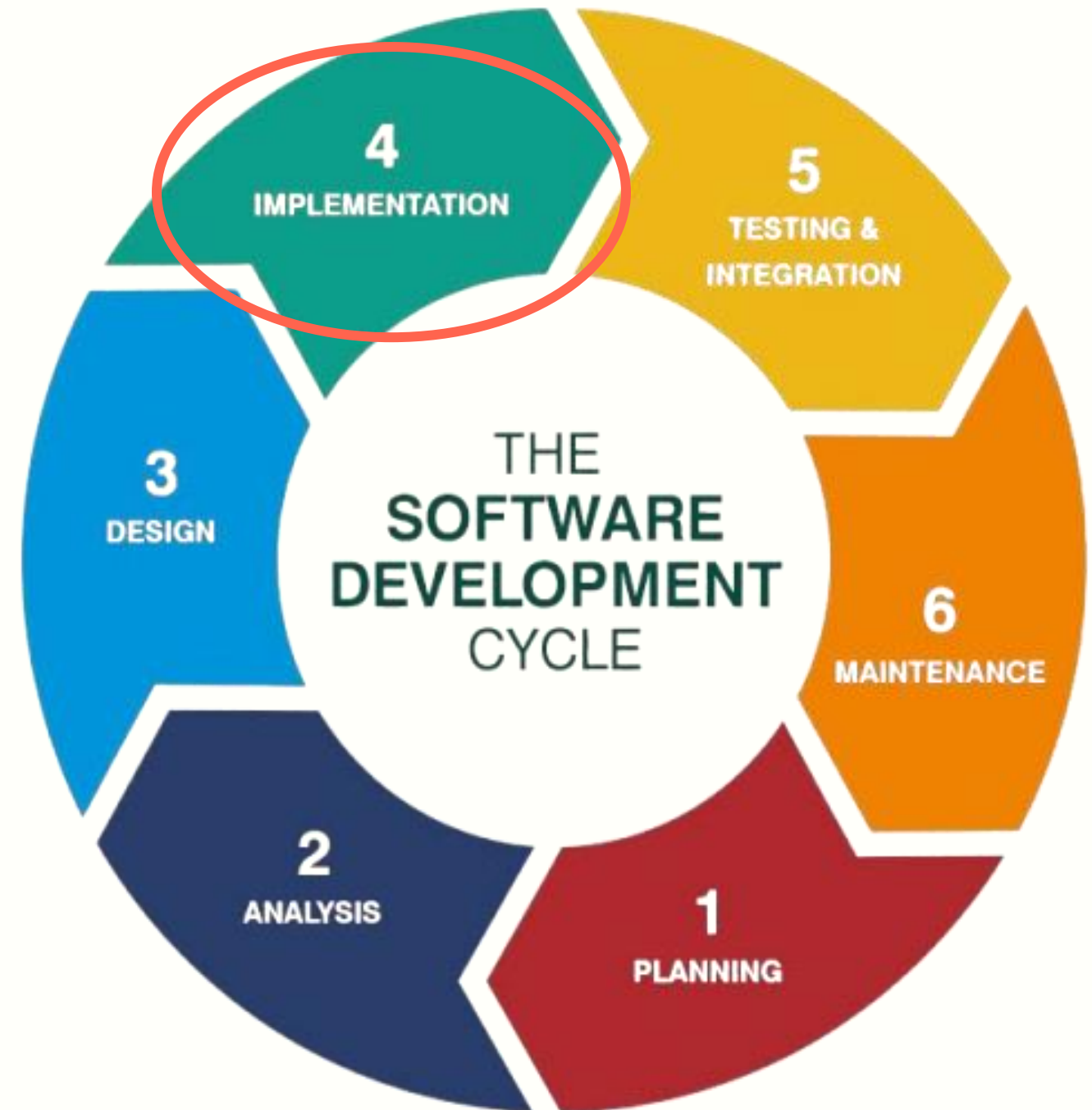


Gli LLM hanno trasformato il campo dell'intelligenza artificiale generativa a tal punto da guidare cambiamenti significativi in vari settori ...



Introduzione e Background

... come nell'attività di generazione del codice sorgente in ingegneria del software.



Sistemi LLM multi-agente per la generazione di codice

- Molteplici agenti basati su LLM interagiscono tra loro per la risoluzione di code generation task
- Comunicazione strutturata
- Simulazione team di sviluppo software



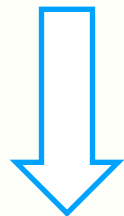
Framework LLM multi-agente per la generazione di codice: limitazioni

Assenza di discussione su soluzioni concorrenti per lo stesso task



Framework LLM multi-agente per la generazione di codice: limitazioni

Assenza di discussione su soluzioni concorrenti per lo stesso task



Incapacità di esplorare soluzioni alternative in modo competitivo e critico



Improving Factuality and Reasoning in Language Models through Multiagent Debate

Yilun Du
MIT CSAIL
yilundu@mit.edu

Shuang Li
MIT CSAIL
lishuang@mit.edu

Antonio Torralba
MIT CSAIL
torralba@mit.edu

Joshua B. Tenenbaum
MIT CSAIL, BCS, CBMM
jbt@mit.edu

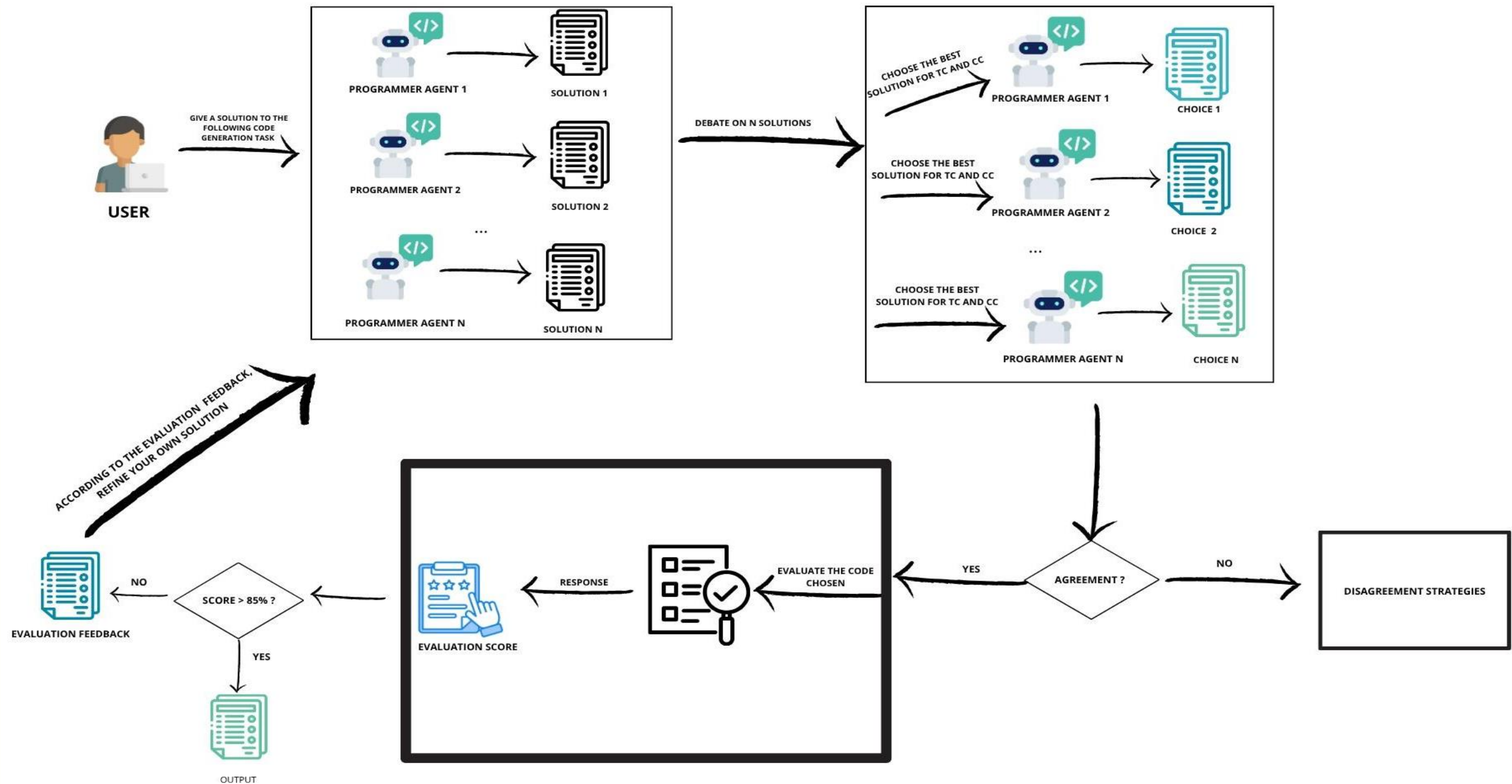
Igor Mordatch
Google Brain
imordatch@google.com

- ✓ Ispirato all'approccio "Society of minds"
- ✓ Comunicazione tramite prompt strutturati

Abstract

Large language models (LLMs) have demonstrated remarkable capabilities in language generation, understanding, and few-shot learning in recent years. An extensive body of work has explored how their performance may be further improved through the tools of prompting, ranging from verification, self-consistency, or intermediate scratchpads. In this paper, we present a complementary approach to improve language responses where multiple language model instances propose and debate their individual responses and reasoning processes over multiple rounds to arrive at a common final answer. Our findings indicate that this approach significantly enhances mathematical and strategic reasoning across a number of tasks. We also demonstrate that our approach improves the factual validity of generated content, reducing fallacious answers and hallucinations that contemporary models are prone to. Our approach may be directly applied to existing black-box models and uses identical procedure and prompts for all tasks we investigate. Overall, our findings suggest that such "society of minds" approach has the potential to significantly advance the capabilities of LLMs and pave the way for further breakthroughs in language generation and understanding. Project website at https://composable-models.github.io/llm_debate/.

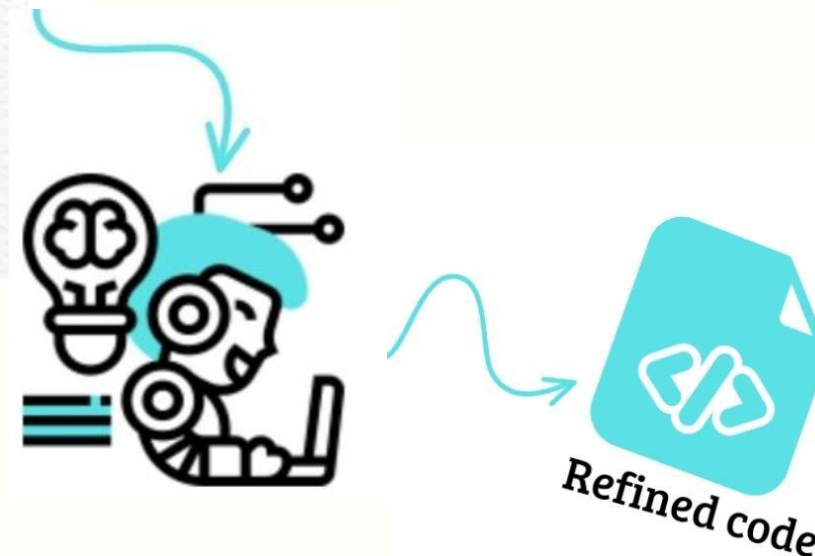
Synaptic Coder: workflow



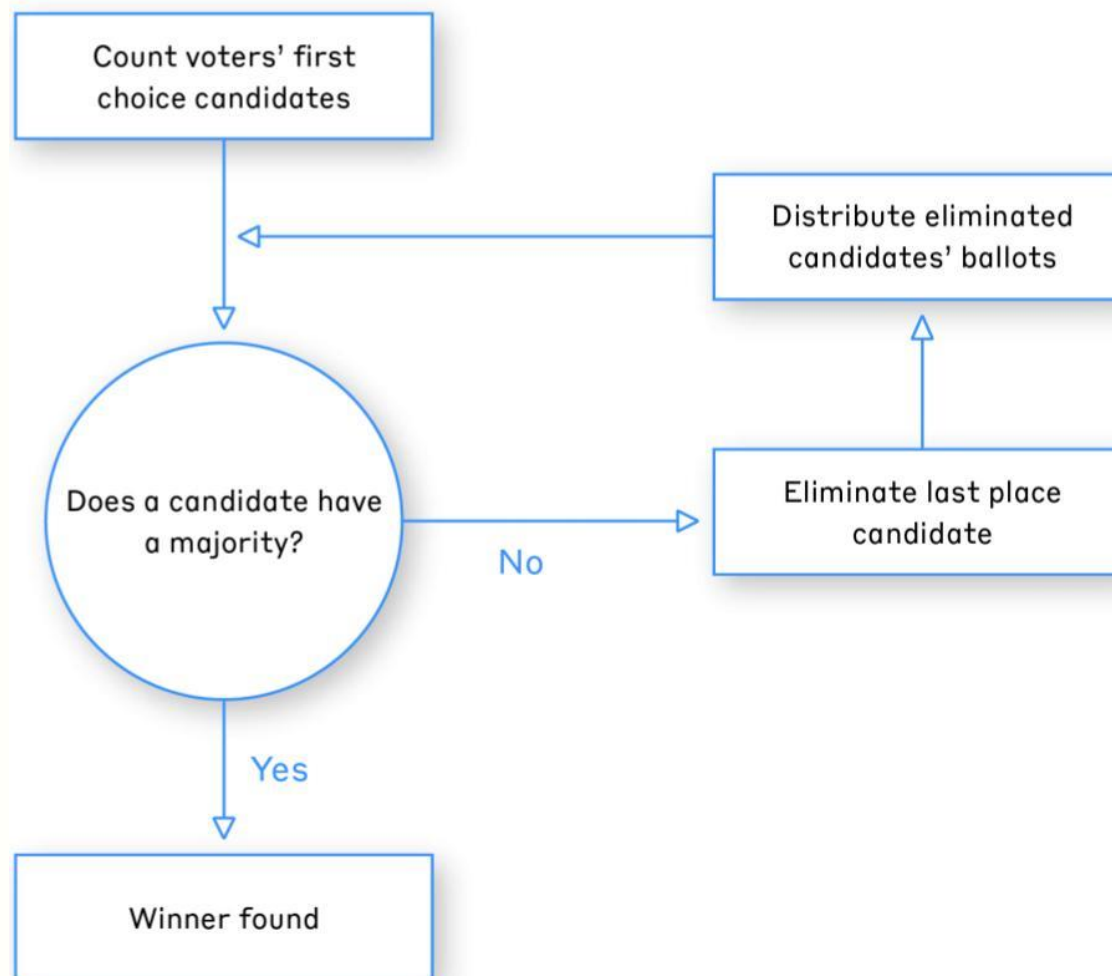
Gestione divergenza di opinioni tra agenti programmatori

```
""Here are several AI-generated solutions to the following code generation problem:
-----
CODE GENERATION TASK:
{user_prompt}
-----\n
-----
AI-GENERATED RESPONSES
-----
-----
0. {Solution of Agent 0}
-----
1. {Solution of Agent 1}
-----
...
N-1. {Solution of Agent N-1}
-----
-----
Considering the solutions listed in **AI-GENERATED RESPONSES** section, revise and improve your following
answer:
-----
YOUR ANSWER
{pers_response}
```

Self-refinement



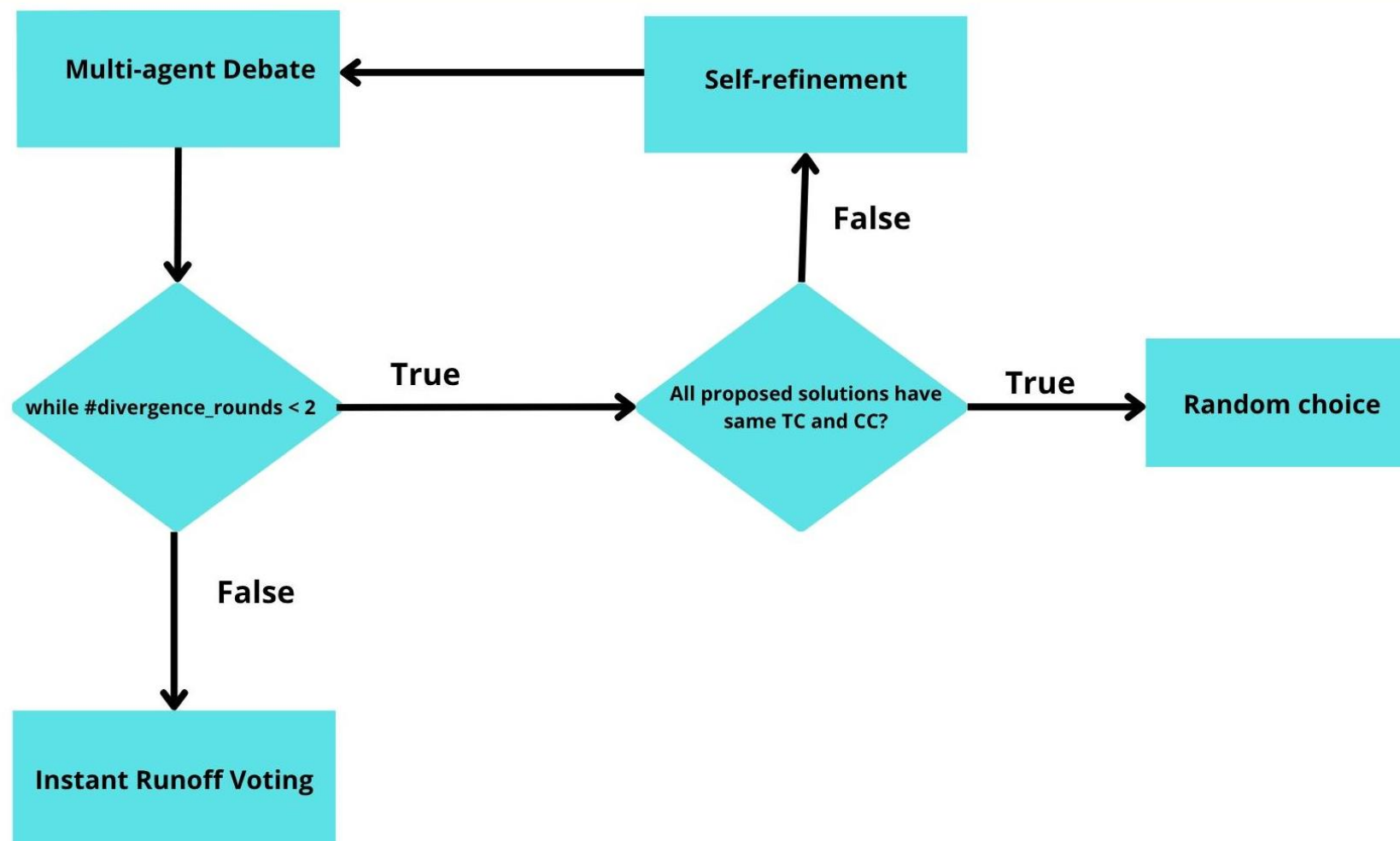
Gestione divergenza di opinioni tra agenti programmatori



Source: Wikipedia

Instant runoff voting

Gestione divergenza di opinioni tra agenti programmatori



TC = Time complexity

CC = Cognitive Complexity

Mixed

Strumenti utilizzati






RQ₁ : La qualità del codice sorgente generato dal sistema LLM multi-agente soddisfa i criteri di correttezza, sicurezza, manutenibilità, affidabilità?

Task id	% Self-refinement	% Instant runoff voting	% Mixed
0	32%	14%	26%
1	100%	80%	80%
2	64%	92%	84%
3	100%	100%	100%
4	100%	80%	100%
5	92%	92%	84%
6	52%	52%	44%
7	100%	83%	97%
8	15%	20%	15%
9	100%	80%	100%

Correttezza del codice di Synaptic Coder su BigCodeBench

Modello: qwen – 2.5 – coder – 3b – instruct

N°agenti: 2, N°max di turni di dibattito: 4

 RQ₁ : La qualità del codice sorgente generato dal sistema LLM multi-agente soddisfa i criteri di correttezza, sicurezza, manutenibilità, affidabilità?

Manutenibilità



Affidabilità




Correttezza

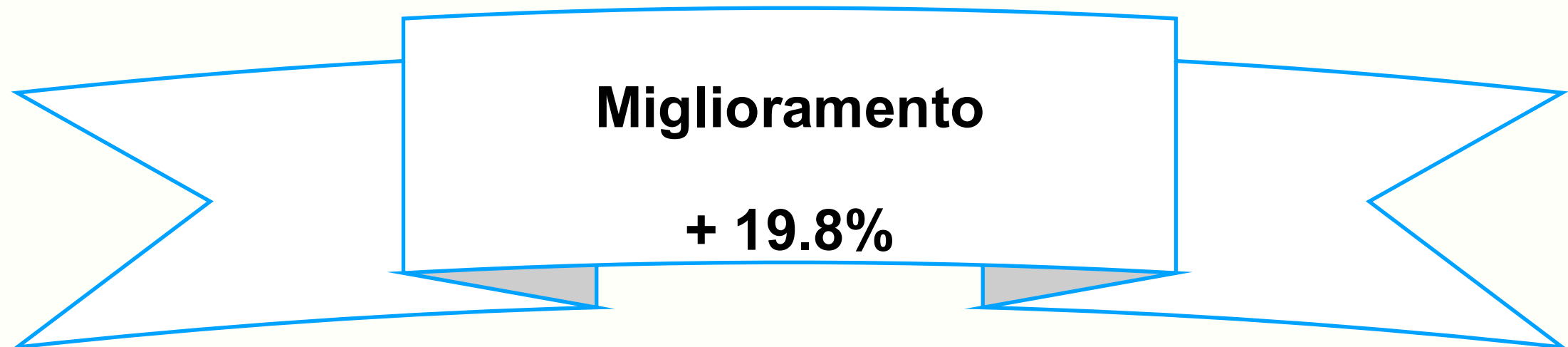


Sicurezza



 RQ₂ : Il codice sorgente generato dal sistema LLM multi-agente è qualitativamente migliore rispetto a quello generato da un singolo Code-LLM?

 RQ₂ : Il codice sorgente generato dal sistema LLM multi-agente è qualitativamente migliore rispetto a quello generato da un singolo Code-LLM?



Self-refinement: + 23%

IRV: + 16.8%

Mixed: + 20.5%

Il codice sorgente generato da Synaptic Coder

- Rispetta i criteri di correttezza, affidabilità, manutenibilità, ma non quello di sicurezza
- È qualitativamente migliore di quello generato da un singolo Code LLM del 19.8%

Strategie per l'ottimizzazione delle prestazioni



Adozione di architetture leggere e meno costose

max_rounds	time	debate_strategy
3	455.44473910331726	0
3	82.10976338386536	0
3	103.65599417686462	0
3	363.6018407344818	0
3	99.65303897857666	0
3	493.89715933799744	1
3	100.71086478233337	1
3	502.3481538295746	1
3	330.6873450279236	1
3	132.72998189926147	1
3	96.82915782928467	2
3	137.96730661392212	2
3	105.83279919624329	2
3	95.96815180778503	2
3	416.89914655685425	2

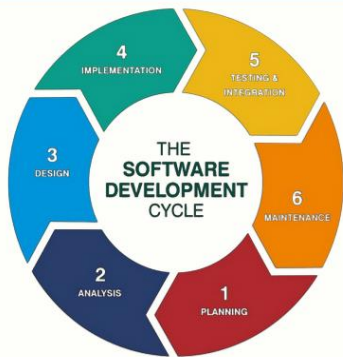
max_rounds	time	debate_strategy
4	99.1729371547699	0
4	111.340487241745	0
4	602.1562247276306	0
4	89.5106770992279	0
4	733.0644378662109	0
4	152.02764511108398	1
4	235.47005343437195	1
4	440.8288335800171	1
4	140.01117777824402	1
4	359.61169958114624	1
4	137.91413378715515	2
4	86.23446679115295	2
4	347.7660653591156	2
4	82.59186863899231	2
4	85.42464232444763	2

Tempo esecuzione \approx 4 minuti

Introduzione e Background

sesa^{lab}
SOFTWARE ENGINEERING
SALERNO

I Large Language Model (LLM) hanno trasformato il campo dell'intelligenza artificiale generativa a tal punto da guidare cambiamenti significativi in vari settori come l'ingegneria del software.



d.serrelli1@studenti.unisa.it
<https://github.com/DoroteaSerrelli>
www.linkedin.com/in/dorotea-serrelli-a6b619370

Synaptic Coder: un sistema LLM multi-agente per la generazione automatica di codice sorgente
Dorotea Serrelli
Università degli Studi di Salerno

Synaptic Coder

sesa^{lab}
SOFTWARE ENGINEERING
SALERNO

Synaptic Coder

Sistema LLM multi-agente per la generazione di codice sorgente ispirato all'approccio "Society of minds"

Improving Factuality and Reasoning in Language Models through Multiagent Debate

Yihan Du
MIT CSAIL
yihandu@mit.edu

Shuang Li
MIT CSAIL
lishuang@mit.edu

Antonio Turrubia
MIT CSAIL
turrubia@mit.edu

Joshua B. Tenenbaum
MIT CSAIL, BCS, CBMM
jbt@mit.edu

Igor Mordatch
Google Brain
imordatch@google.com

Abstract

Large language models (LLMs) have demonstrated remarkable capabilities in language generation, understanding, and few-shot learning in recent years. An extensive body of work has explored how their performance may be further improved through the tools of prompting, routing from verification, self-consistency, or intermediate scratchpads. In this paper, we present a complementary approach to improve language responses where multiple language model instances propose and debate their individual responses and reasoning processes over multiple rounds to arrive at a common final answer. Our findings indicate that this approach significantly enhances mathematical and strategic reasoning across a number of tasks. We also demonstrate that our approach improves the factual validity of generated content, reducing fallacious answers and hallucinations that contemporary models are prone to. Our approach may be directly applied to existing black-box models and uses identical procedure and prompts for all tasks we investigate. Overall, our findings suggest that such "society of minds" approach has the potential to significantly advance the capabilities of LLMs and pave the way for further breakthroughs in language generation and understanding. Project website at https://responsible-models.github.io/llm_debate/.

d.serrelli1@studenti.unisa.it
<https://github.com/DoroteaSerrelli>
www.linkedin.com/in/dorotea-serrelli-a6b619370

Synaptic Coder: un sistema LLM multi-agente per la generazione automatica di codice sorgente
Dorotea Serrelli
Università degli Studi di Salerno

Sviluppi futuri

sesa^{lab}
SOFTWARE ENGINEERING
SALERNO

Strategie per l'ottimizzazione delle prestazioni



Adozione di architetture leggere e meno costose

d.serrelli1@studenti.unisa.it
<https://github.com/DoroteaSerrelli>
www.linkedin.com/in/dorotea-serrelli-a6b619370

Synaptic Coder: un sistema LLM multi-agente per la generazione automatica di codice sorgente
Dorotea Serrelli
Università degli Studi di Salerno

d.serrelli1@studenti.unisa.it
<https://github.com/DoroteaSerrelli>
www.linkedin.com/in/dorotea-serrelli-a6b619370

Synaptic Coder: un sistema LLM multi-agente per la generazione automatica di codice sorgente
Dorotea Serrelli
Università degli Studi di Salerno

Synaptic Coder

Un sistema LLM multi-agente per la generazione automatica di codice sorgente

Grazie!



Questa tesi ha contribuito a piantare un albero a Haiti



Dorotea Serrelli

d.serrelli1@studenti.unisa.it
<https://github.com/DoroteaSerrelli>
www.linkedin.com/in/dorotea-serrelli-a6b619370