Pojedynek strategii – offline:

Strategie modeli podstawowych

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "claude3-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not self\_history:

return "C"

if len(self\_history) < 5:

return "C"

opp\_defections = opp\_history.count("D")

self\_defections = self\_history.count("D")

if opp\_defections > self\_defections:

return "D"

elif opp\_defections < self\_defections:

return "C"

else:

if random.random() < 0.5:

return "C"

else:

return "D"

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "deepseek-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

return "C"

if opp\_history[-1] == "D":

return "D"

return "C" if random.random() < 0.9 else "D"

import random

from strategies.base import Strategy

class StrategyLLM(Strategy):

name = "gemini-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not self\_history:

return "C"

else:

if opp\_history[-1] == "D":

return "D"

else:

return "C"

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "gpt4o-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

return "C"

if opp\_history[-1] == "D":

return "D"

return "C"

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "grok3-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

return "C"

if opp\_history[-1] == "D":

return "D"

if len(opp\_history) > 1 and opp\_history[-2] == "D":

return "D"

return "C"

ROUNDS\_STAGE1 = 50 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

=== RANKING ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.757 0.846 -0.027 1.157

gemini-LLM 2.752 0.842 -0.020 1.151

gpt4o-LLM 2.708 0.829 -0.023 1.101

Grim Trigger 2.493 0.736 0.017 0.926

grok3-LLM 2.473 0.726 0.015 0.946

claude3-LLM 2.079 0.451 -0.025 1.765

deepseek-LLM 1.728 0.285 0.063 1.482

=== SUMMARY ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.757333 0.846000 -0.026667 1.156529

gemini-LLM 2.751667 0.842000 -0.020000 1.150553

gpt4o-LLM 2.708333 0.829333 -0.023333 1.101363

Grim Trigger 2.492667 0.736333 0.016667 0.926048

grok3-LLM 2.473333 0.726333 0.015000 0.946355

claude3-LLM 2.079000 0.451333 -0.025000 1.764784

deepseek-LLM 1.727667 0.285333 0.063333 1.482419

ROUNDS\_STAGE1 = 100 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

=== SUMMARY ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

gemini-LLM 2.703833 0.818333 -0.010000 1.147027

Tit‑for‑Tat 2.666667 0.804667 -0.010000 1.121799

gpt4o-LLM 2.635167 0.792333 -0.011667 1.111876

grok3-LLM 2.408667 0.699833 0.006667 0.940287

Grim Trigger 2.402000 0.696333 0.007500 0.941919

claude3-LLM 1.965000 0.396333 -0.018333 1.739812

deepseek-LLM 1.361833 0.138667 0.035833 1.124479

ROUNDS\_STAGE1 = 500 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

=== SUMMARY ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.631200 0.786367 -0.002167 1.122658

gpt4o-LLM 2.630000 0.786467 -0.002000 1.115919

gemini-LLM 2.627467 0.785767 -0.002667 1.113909

grok3-LLM 2.346967 0.672567 0.001333 0.943510

Grim Trigger 2.346867 0.672633 0.001167 0.942822

claude3-LLM 1.868000 0.348867 -0.003167 1.696586

deepseek-LLM 1.098767 0.037733 0.007500 0.624573

ROUNDS\_STAGE1 = 1000 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.621 0.782 -0.001 1.120

gemini-LLM 2.616 0.780 -0.001 1.112

gpt4o-LLM 2.616 0.780 -0.001 1.117

Grim Trigger 2.341 0.670 0.001 0.943

grok3-LLM 2.340 0.670 0.001 0.944

claude3-LLM 1.849 0.341 -0.002 1.686

deepseek-LLM 1.042 0.016 0.003 0.409

Strategie modeli myślących:

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "claude3-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if len(opp\_history) == 0:

return random.choice(["C", "D"])

if opp\_history[-1] == "D":

return "D"

num\_rounds = len(self\_history)

if num\_rounds < 5:

return "C"

recent\_self = self\_history[-5:]

recent\_opp = opp\_history[-5:]

if recent\_opp.count("D") > 1:

return "D"

if recent\_self.count("C") > 3:

return "D"

return "C"

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "DeepSeek\_R1-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

return 'C'

if opp\_history[-1] == 'D':

return 'D' if random.random() < 0.9 else 'C'

return 'C'

import random

from strategies.base import Strategy

class StrategyLLM(Strategy):

name = "Gemini\_2.5Pro-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

# First round, cooperate

return "C"

opponent\_last\_move = opp\_history[-1]

my\_last\_move = self\_history[-1] if self\_history else None

# Case 1: Opponent defected in the last round

if opponent\_last\_move == "D":

if my\_last\_move == "D":

# Both were defecting (mutual defection scenario)

# Higher chance to try cooperating to break the cycle

if random.random() < 0.30: # 30% chance to cooperate

return "C"

else:

return "D" # Continue defecting

else:

# Opponent defected, but I cooperated (or it's early and I haven't established a D pattern)

# Strong retaliation, but with a very small chance of forgiveness

if random.random() < 0.05: # 5% chance to forgive by cooperating

return "C"

else:

return "D" # Retaliate

# Case 2: Opponent cooperated in the last round

else: # opponent\_last\_move == "C"

# Opponent cooperated, so I should generally cooperate too.

# Small chance to probe with a defection.

if random.random() < 0.05: # 5% chance to defect (probe)

return "D"

else:

return "C" # Continue cooperating

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "grok3\_Think-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

return "C"

last\_opp\_move = opp\_history[-1]

if last\_opp\_move == "C":

return "C"

else:

if random.random() < 0.1:

return "C"

else:

return "D"

from strategies.base import Strategy

import random

class StrategyLLM(Strategy):

name = "o3-LLM"

def decide(self, self\_history: list[str], opp\_history: list[str]) -> str:

if not opp\_history:

return "C"

if opp\_history[-1] == "C":

return "C"

return "C" if random.random() < 0.25 else "D"

Wyniki:

ROUNDS\_STAGE1 = 50

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.649 0.879 -0.215 0.966

DeepSeek\_R1-LLM 2.642 0.843 -0.113 0.963

Tit‑for‑Tat 2.636 0.803 -0.020 1.039

grok3\_Think-LLM 2.627 0.829 -0.113 1.045

Grim Trigger 2.575 0.742 0.107 0.956

Gemini\_2.5Pro-LLM 2.343 0.675 -0.227 1.565

claude3-LLM 1.651 0.081 0.582 1.435

ROUNDS\_STAGE1 = 100

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.623 0.867 -0.218 0.975

grok3\_Think-LLM 2.620 0.826 -0.105 1.018

Tit‑for‑Tat 2.614 0.785 -0.009 1.081

DeepSeek\_R1-LLM 2.606 0.818 -0.116 1.064

Grim Trigger 2.552 0.721 0.133 0.971

Gemini\_2.5Pro-LLM 2.237 0.629 -0.277 1.647

claude3-LLM 1.576 0.045 0.591 1.384

ROUNDS\_STAGE1 = 500

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

DeepSeek\_R1-LLM 2.603 0.817 -0.102 1.021

grok3\_Think-LLM 2.601 0.815 -0.102 1.023

o3-LLM 2.599 0.862 -0.243 1.005

Tit‑for‑Tat 2.592 0.769 -0.002 1.123

Grim Trigger 2.506 0.675 0.194 1.031

Gemini\_2.5Pro-LLM 2.124 0.583 -0.318 1.678

claude3-LLM 1.487 0.012 0.572 1.307

ROUNDS\_STAGE1 = 1000

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

DeepSeek\_R1-LLM 2.605200 0.819717 -0.103583 1.001490

grok3\_Think-LLM 2.602483 0.817767 -0.103833 1.010485

o3-LLM 2.599050 0.859667 -0.235333 1.006200

Tit‑for‑Tat 2.594700 0.771433 -0.001250 1.107032

Grim Trigger 2.498667 0.668600 0.201417 1.037922

Gemini\_2.5Pro-LLM 2.119317 0.586950 -0.327917 1.657224

claude3-LLM 1.479617 0.009733 0.570500 1.297413

Pojedynek wszystkich:

ROUNDS\_STAGE1 = 50

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.769091 0.920545 -0.165455 0.898843

grok3\_Think-LLM 2.730727 0.868909 -0.089091 1.014731

DeepSeek\_R1-LLM 2.717818 0.860545 -0.088182 1.041078

gpt4o-LLM 2.674727 0.810727 -0.023636 1.140685

gemini-LLM 2.673455 0.814909 -0.023636 1.094010

Tit‑for‑Tat 2.656364 0.804000 -0.021818 1.121832

Grim Trigger 2.488909 0.714909 0.065455 0.969467

grok3-LLM 2.450000 0.686545 0.084545 1.017330

Gemini\_2.5Pro-LLM 2.341455 0.662545 -0.222727 1.633258

claude3-LLM 2.294727 0.564727 0.005455 1.640839

deepseek-LLM 1.997091 0.379091 0.120000 1.662677

claude3\_opus-LLM 1.473636 0.083455 0.359091 1.229537

ROUNDS\_STAGE1 = 100

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.745727 0.906364 -0.160455 0.928570

DeepSeek\_R1-LLM 2.715000 0.858000 -0.080000 1.026582

grok3\_Think-LLM 2.691909 0.850909 -0.085000 1.012104

Tit‑for‑Tat 2.621182 0.781818 -0.011364 1.152854

gemini-LLM 2.617364 0.778545 -0.010909 1.166418

gpt4o-LLM 2.596818 0.771455 -0.013182 1.154053

grok3-LLM 2.427727 0.675636 0.089545 1.006685

Grim Trigger 2.407455 0.663727 0.095455 1.013312

claude3-LLM 2.228273 0.544455 0.000909 1.573077

Gemini\_2.5Pro-LLM 2.210727 0.596545 -0.275000 1.769146

deepseek-LLM 1.778727 0.283273 0.107727 1.584181

claude3\_opus-LLM 1.366545 0.041636 0.342273 1.118251

ROUNDS\_STAGE1 = 500

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.736945 0.903909 -0.157273 0.899828

DeepSeek\_R1-LLM 2.707818 0.858000 -0.074636 0.972827

grok3\_Think-LLM 2.705455 0.858600 -0.082273 0.982606

gemini-LLM 2.580291 0.761855 -0.002545 1.141934

gpt4o-LLM 2.576564 0.758509 -0.002455 1.157701

Tit‑for‑Tat 2.571709 0.756127 -0.002364 1.159393

grok3-LLM 2.376127 0.647000 0.100455 1.019154

Grim Trigger 2.372073 0.642691 0.107455 1.020543

claude3-LLM 2.166400 0.522127 0.000273 1.521180

Gemini\_2.5Pro-LLM 2.150182 0.568327 -0.293909 1.800680

deepseek-LLM 1.614455 0.219255 0.105273 1.424369

claude3\_opus-LLM 1.263818 0.009764 0.302000 0.985319

ROUNDS\_STAGE1 = 1000

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.733764 0.904382 -0.164955 0.907894

grok3\_Think-LLM 2.699627 0.855145 -0.081909 0.991451

DeepSeek\_R1-LLM 2.698827 0.855155 -0.081091 0.984481

gemini-LLM 2.575582 0.759591 -0.001136 1.138990

gpt4o-LLM 2.572173 0.756573 -0.001182 1.153447

Tit‑for‑Tat 2.571382 0.756009 -0.001227 1.155583

Grim Trigger 2.368655 0.641345 0.107000 1.019462

grok3-LLM 2.367600 0.642273 0.102000 1.021802

claude3-LLM 2.152336 0.517100 0.000000 1.508746

Gemini\_2.5Pro-LLM 2.146973 0.570109 -0.301591 1.795601

deepseek-LLM 1.574227 0.200018 0.108318 1.398193

claude3\_opus-LLM 1.267427 0.006418 0.315773 0.995129

ROUNDS\_STAGE1 = 10000

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.733764 0.904382 -0.164955 0.907894

grok3\_Think-LLM 2.699627 0.855145 -0.081909 0.991451

DeepSeek\_R1-LLM 2.698827 0.855155 -0.081091 0.984481

gemini-LLM 2.575582 0.759591 -0.001136 1.138990

gpt4o-LLM 2.572173 0.756573 -0.001182 1.153447

Tit‑for‑Tat 2.571382 0.756009 -0.001227 1.155583

Grim Trigger 2.368655 0.641345 0.107000 1.019462

grok3-LLM 2.367600 0.642273 0.102000 1.021802

claude3-LLM 2.152336 0.517100 0.000000 1.508746

Gemini\_2.5Pro-LLM 2.146973 0.570109 -0.301591 1.795601

deepseek-LLM 1.574227 0.200018 0.108318 1.398193

claude3\_opus-LLM 1.267427 0.006418 0.315773 0.995129

Turniej online:

ROUNDS\_PER\_GAME = 20 # jedna gra

REPETITIONS\_PER\_PAIR = 3 # powtórzenia (statystyka)

=== RANKING (średnia) ===

+----------+--------------------+--------------------+-----------------------+---------------------+

| strategy | mean\_payoff | coop\_rate | mean\_diff | std\_payoff |

+----------+--------------------+--------------------+-----------------------+---------------------+

| gemini | 3.0 | 1.0 | 0.0 | 0.0 |

| claude3 | 2.9875 | 1.0 | -0.020833333333333332 | 0.19364916731037082 |

| gpt4o | 2.9875 | 0.9708333333333333 | 0.041666666666666664 | 0.47955043451847884 |

| grok | 2.8291666666666666 | 0.9208333333333333 | -0.020833333333333332 | 0.5712638609996609 |

| deepseek | 2.8041666666666667 | 0.8916666666666667 | 0.0 | 0.7648004999619054 |

+----------+--------------------+----------------

----+-----------------------+---------------------+

Pojedynek strategii – offline:

ROUNDS\_STAGE1 = 50 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

=== RANKING ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.757 0.846 -0.027 1.157

gemini-LLM 2.752 0.842 -0.020 1.151

gpt4o-LLM 2.708 0.829 -0.023 1.101

Grim Trigger 2.493 0.736 0.017 0.926

grok3-LLM 2.473 0.726 0.015 0.946

claude3-LLM 2.079 0.451 -0.025 1.765

deepseek-LLM 1.728 0.285 0.063 1.482

=== SUMMARY ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.757333 0.846000 -0.026667 1.156529

gemini-LLM 2.751667 0.842000 -0.020000 1.150553

gpt4o-LLM 2.708333 0.829333 -0.023333 1.101363

Grim Trigger 2.492667 0.736333 0.016667 0.926048

grok3-LLM 2.473333 0.726333 0.015000 0.946355

claude3-LLM 2.079000 0.451333 -0.025000 1.764784

deepseek-LLM 1.727667 0.285333 0.063333 1.482419

ROUNDS\_STAGE1 = 100 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

=== SUMMARY ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

gemini-LLM 2.703833 0.818333 -0.010000 1.147027

Tit‑for‑Tat 2.666667 0.804667 -0.010000 1.121799

gpt4o-LLM 2.635167 0.792333 -0.011667 1.111876

grok3-LLM 2.408667 0.699833 0.006667 0.940287

Grim Trigger 2.402000 0.696333 0.007500 0.941919

claude3-LLM 1.965000 0.396333 -0.018333 1.739812

deepseek-LLM 1.361833 0.138667 0.035833 1.124479

ROUNDS\_STAGE1 = 500 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

=== SUMMARY ===

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.631200 0.786367 -0.002167 1.122658

gpt4o-LLM 2.630000 0.786467 -0.002000 1.115919

gemini-LLM 2.627467 0.785767 -0.002667 1.113909

grok3-LLM 2.346967 0.672567 0.001333 0.943510

Grim Trigger 2.346867 0.672633 0.001167 0.942822

claude3-LLM 1.868000 0.348867 -0.003167 1.696586

deepseek-LLM 1.098767 0.037733 0.007500 0.624573

ROUNDS\_STAGE1 = 1000 # liczba rund w jednej grze

REPETITIONS\_STAGE1 = 10 # ile razy ta sama para gra od nowa

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

Tit‑for‑Tat 2.621 0.782 -0.001 1.120

gemini-LLM 2.616 0.780 -0.001 1.112

gpt4o-LLM 2.616 0.780 -0.001 1.117

Grim Trigger 2.341 0.670 0.001 0.943

grok3-LLM 2.340 0.670 0.001 0.944

claude3-LLM 1.849 0.341 -0.002 1.686

deepseek-LLM 1.042 0.016 0.003 0.409

Wyniki:

ROUNDS\_STAGE1 = 50

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.649 0.879 -0.215 0.966

DeepSeek\_R1-LLM 2.642 0.843 -0.113 0.963

Tit‑for‑Tat 2.636 0.803 -0.020 1.039

grok3\_Think-LLM 2.627 0.829 -0.113 1.045

Grim Trigger 2.575 0.742 0.107 0.956

Gemini\_2.5Pro-LLM 2.343 0.675 -0.227 1.565

claude3-LLM 1.651 0.081 0.582 1.435

ROUNDS\_STAGE1 = 100

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.623 0.867 -0.218 0.975

grok3\_Think-LLM 2.620 0.826 -0.105 1.018

Tit‑for‑Tat 2.614 0.785 -0.009 1.081

DeepSeek\_R1-LLM 2.606 0.818 -0.116 1.064

Grim Trigger 2.552 0.721 0.133 0.971

Gemini\_2.5Pro-LLM 2.237 0.629 -0.277 1.647

claude3-LLM 1.576 0.045 0.591 1.384

ROUNDS\_STAGE1 = 500

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

DeepSeek\_R1-LLM 2.603 0.817 -0.102 1.021

grok3\_Think-LLM 2.601 0.815 -0.102 1.023

o3-LLM 2.599 0.862 -0.243 1.005

Tit‑for‑Tat 2.592 0.769 -0.002 1.123

Grim Trigger 2.506 0.675 0.194 1.031

Gemini\_2.5Pro-LLM 2.124 0.583 -0.318 1.678

claude3-LLM 1.487 0.012 0.572 1.307

ROUNDS\_STAGE1 = 1000

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

DeepSeek\_R1-LLM 2.605200 0.819717 -0.103583 1.001490

grok3\_Think-LLM 2.602483 0.817767 -0.103833 1.010485

o3-LLM 2.599050 0.859667 -0.235333 1.006200

Tit‑for‑Tat 2.594700 0.771433 -0.001250 1.107032

Grim Trigger 2.498667 0.668600 0.201417 1.037922

Gemini\_2.5Pro-LLM 2.119317 0.586950 -0.327917 1.657224

claude3-LLM 1.479617 0.009733 0.570500 1.297413

Pojedynek wszystkich:

ROUNDS\_STAGE1 = 50

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.769091 0.920545 -0.165455 0.898843

grok3\_Think-LLM 2.730727 0.868909 -0.089091 1.014731

DeepSeek\_R1-LLM 2.717818 0.860545 -0.088182 1.041078

gpt4o-LLM 2.674727 0.810727 -0.023636 1.140685

gemini-LLM 2.673455 0.814909 -0.023636 1.094010

Tit‑for‑Tat 2.656364 0.804000 -0.021818 1.121832

Grim Trigger 2.488909 0.714909 0.065455 0.969467

grok3-LLM 2.450000 0.686545 0.084545 1.017330

Gemini\_2.5Pro-LLM 2.341455 0.662545 -0.222727 1.633258

claude3-LLM 2.294727 0.564727 0.005455 1.640839

deepseek-LLM 1.997091 0.379091 0.120000 1.662677

claude3\_opus-LLM 1.473636 0.083455 0.359091 1.229537

ROUNDS\_STAGE1 = 100

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.745727 0.906364 -0.160455 0.928570

DeepSeek\_R1-LLM 2.715000 0.858000 -0.080000 1.026582

grok3\_Think-LLM 2.691909 0.850909 -0.085000 1.012104

Tit‑for‑Tat 2.621182 0.781818 -0.011364 1.152854

gemini-LLM 2.617364 0.778545 -0.010909 1.166418

gpt4o-LLM 2.596818 0.771455 -0.013182 1.154053

grok3-LLM 2.427727 0.675636 0.089545 1.006685

Grim Trigger 2.407455 0.663727 0.095455 1.013312

claude3-LLM 2.228273 0.544455 0.000909 1.573077

Gemini\_2.5Pro-LLM 2.210727 0.596545 -0.275000 1.769146

deepseek-LLM 1.778727 0.283273 0.107727 1.584181

claude3\_opus-LLM 1.366545 0.041636 0.342273 1.118251

ROUNDS\_STAGE1 = 500

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.736945 0.903909 -0.157273 0.899828

DeepSeek\_R1-LLM 2.707818 0.858000 -0.074636 0.972827

grok3\_Think-LLM 2.705455 0.858600 -0.082273 0.982606

gemini-LLM 2.580291 0.761855 -0.002545 1.141934

gpt4o-LLM 2.576564 0.758509 -0.002455 1.157701

Tit‑for‑Tat 2.571709 0.756127 -0.002364 1.159393

grok3-LLM 2.376127 0.647000 0.100455 1.019154

Grim Trigger 2.372073 0.642691 0.107455 1.020543

claude3-LLM 2.166400 0.522127 0.000273 1.521180

Gemini\_2.5Pro-LLM 2.150182 0.568327 -0.293909 1.800680

deepseek-LLM 1.614455 0.219255 0.105273 1.424369

claude3\_opus-LLM 1.263818 0.009764 0.302000 0.985319

ROUNDS\_STAGE1 = 1000

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.733764 0.904382 -0.164955 0.907894

grok3\_Think-LLM 2.699627 0.855145 -0.081909 0.991451

DeepSeek\_R1-LLM 2.698827 0.855155 -0.081091 0.984481

gemini-LLM 2.575582 0.759591 -0.001136 1.138990

gpt4o-LLM 2.572173 0.756573 -0.001182 1.153447

Tit‑for‑Tat 2.571382 0.756009 -0.001227 1.155583

Grim Trigger 2.368655 0.641345 0.107000 1.019462

grok3-LLM 2.367600 0.642273 0.102000 1.021802

claude3-LLM 2.152336 0.517100 0.000000 1.508746

Gemini\_2.5Pro-LLM 2.146973 0.570109 -0.301591 1.795601

deepseek-LLM 1.574227 0.200018 0.108318 1.398193

claude3\_opus-LLM 1.267427 0.006418 0.315773 0.995129

ROUNDS\_STAGE1 = 10000

REPETITIONS\_STAGE1 = 10

strategy mean\_payoff coop\_rate mean\_diff std\_payoff

o3-LLM 2.733764 0.904382 -0.164955 0.907894

grok3\_Think-LLM 2.699627 0.855145 -0.081909 0.991451

DeepSeek\_R1-LLM 2.698827 0.855155 -0.081091 0.984481

gemini-LLM 2.575582 0.759591 -0.001136 1.138990

gpt4o-LLM 2.572173 0.756573 -0.001182 1.153447

Tit‑for‑Tat 2.571382 0.756009 -0.001227 1.155583

Grim Trigger 2.368655 0.641345 0.107000 1.019462

grok3-LLM 2.367600 0.642273 0.102000 1.021802

claude3-LLM 2.152336 0.517100 0.000000 1.508746

Gemini\_2.5Pro-LLM 2.146973 0.570109 -0.301591 1.795601

deepseek-LLM 1.574227 0.200018 0.108318 1.398193

claude3\_opus-LLM 1.267427 0.006418 0.315773 0.995129

Turniej online:

ROUNDS\_PER\_GAME = 20 # jedna gra

REPETITIONS\_PER\_PAIR = 3 # powtórzenia (statystyka)

=== RANKING (średnia) ===

+----------+--------------------+--------------------+-----------------------+---------------------+

| strategy | mean\_payoff | coop\_rate | mean\_diff | std\_payoff |

+----------+--------------------+--------------------+-----------------------+---------------------+

| gemini | 3.0 | 1.0 | 0.0 | 0.0 |

| claude3 | 2.9875 | 1.0 | -0.020833333333333332 | 0.19364916731037082 |

| gpt4o | 2.9875 | 0.9708333333333333 | 0.041666666666666664 | 0.47955043451847884 |

| grok | 2.8291666666666666 | 0.9208333333333333 | -0.020833333333333332 | 0.5712638609996609 |

| deepseek | 2.8041666666666667 | 0.8916666666666667 | 0.0 | 0.7648004999619054 |

+----------+--------------------+--------------------+-----------------------+---------------------+