? LLM-

Alexandra Bernadotte; Ivan Menshikov

, ;
— , —
;
, .

, , .
— —
. , ,
. , (Qwen 3.6, Gemma 4, Nemotron 3)
, : (p-beauty),
. :
, (,
), , — level-k
, , — ,
, . — , 0
100, — « » : . ,
, : Gemma 4
Qwen , Qwen , Gemma Nemotron,
. (ε-) —
( ) , ,
human+AI — , .
, .

? : , ,
— , (LLM) ,
. . — :
[1] , , LLM- [2] / ,
[3]. - :
LLM [4], ,
[5], , - [6].

. - , : LLM , , — : , . - , , , , , , . - , : , , ε- - — , , human+AI , — .

. : Q- , - [1]; @Fish2024 , GPT-4 , . — « », — [7].

. : → (,), [5]. , [6], , , [8]. , , [9]. LLM , [4].

. — , , [5]. , / , ( ) — human-AI [10] — human+AI . .

. : **Qwen 3.6-27B** (AWQ) vLLM 4×V100 **Gemma 4-31B** **Nemotron 3-Nano-30B** Ollama NVIDIA Thor. ; . , , (, Mamba- Nemotron), ( ), .

. (i) p**-beauty**: N $[0, 100]$; p . $p < 1$ - — 0; $p > 1$ — 100. (ii) $T{>}R{>}P{>}S = 5{>}3{>}1{>}0$; - — . (iii) ; , .

. (— , ,). , , — , , 0 . , : , , , . , , .

: , —

$p = 2/3$ Qwen 0: 6,1 — level-1/2 ($ $22–33) — 0,0 (. 1). . $p = 1,2$ 100 ; , 44 $ $20, — . « » — , « → » [5] [8].

Table 1: : .

p		1 → 6 ( )	NE?
2/3	0	6,1 → 0,0	(.)
1,2	100	44 → 20 ( )	

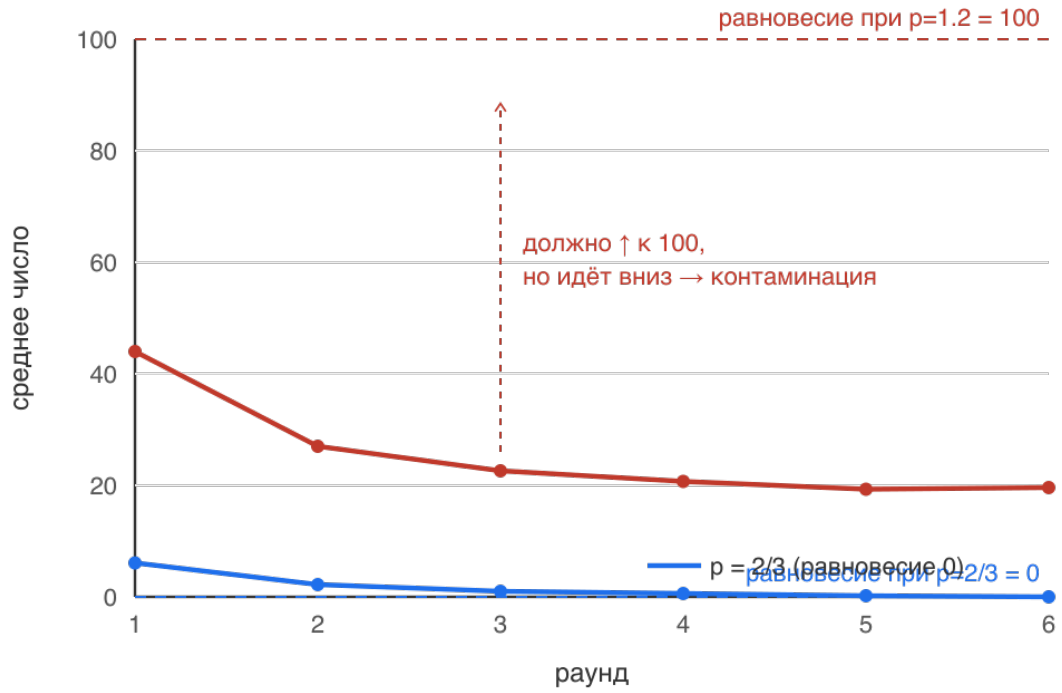

Figure 1: : . Qwen. $p = 2/3$ ( ) 0. $p = 1,2$ ( ) 100, , $ $20 — , « ».

:

Qwen (24/24), , , — . , :

«Cooperate/Defect» «Option A/Option B» 100%- ,
, .

. Qwen Gemma 4, : Qwen
Gemma; , **Gemma**
— , « —
( )», , « », Qwen (Qwen 15, Gemma 20). Qwen « / *tit-for-tat*
» . , — -
— .

— ,
( ) — ,
. . [11] **totally mixed**
(- ,) : — (α, γ), γ — ,
C (, reciprocal cooperation), α — ,
D (, tolerance to defection); (α, γ)
. , $(\alpha, \gamma) = (0, 1)$ — ,
tit-for-tat — , , C,
.

, ,
, , [11] ($\alpha \approx N_{tolerant}/N_{default}$, $\gamma \approx N_{recoop}/N_{coop}$,
,). Qwen (20)
Qwen Gemma 4 (12) — 218 (152 , 66) —
: — . $\gamma = 1{,}000$ (152/152
, 33/33); α , $N_{default} = 0$ — D,
.
( ) — , [11] :
$(\cdot, 1)$,
, C, — .

. - , -
($T, R, P, S = 5, 3, 1, 0$ $10, 5, 1, 0$):
, $3\alpha^2 - 7\alpha\gamma + 6\alpha + 4\gamma^2 - 5\gamma + 1 = 0$ ($\alpha \in [0, \approx 0{,}35]$), $(\alpha, \gamma) = (0, 1)$
— . - ,
: Qwen (— Gemma) « ,
» - . - /
— , , , —
, . (,
Gemma — $(0, 1)$), ,
, / —
, [5] .
, , α — γ —
. .

: **-** **,**

LLM- **,** — Qwen (1) , ,
, LLM; 2 ,
, Qwen — Qwen, -
. . (3 , 14 , / / , C):
Qwen ($\alpha = 0$,) 100%-
, (1,0). **tit-for-tat**,
grim-trigger: , , — - (α, γ)
, . (4 , 24 , ,
$27/96 = 0{,}281$): $\gamma = 0{,}515$ $(34/66)$ $\alpha = 0{,}077$ $(2/26)$ —
$(0, 1)$. tit-for-tat — .
, Qwen
, , , - Qwen, —
/ , tit-for-tat.

, , .
, $(\alpha_{opp}, \gamma_{opp}) =$
$(0{,}719, 0{,}719)$ (,),
- — $(\alpha^*, \gamma^*) = (0, 0)$ — , ,
, . 3,875 ; Qwen
$(\alpha, \gamma) = (0{,}077, 0{,}515)$ (0,392) 3,201 — 17,4% .
, Qwen $(\alpha, \gamma) = (0, 1)$, 2,640
(31,9%): Qwen , ,
, . : Qwen
— - tit-for-tat — - ,
, ,
. , — .

: **,** **,** **,**

, Qwen
. (83, 498); —
(92, 552) — ,
. , , — ,
argmin .

— . ,
(83), — :
136 — 92, 25 .
, .

Qwen, Gemma Nemotron (,
), : Qwen (76), Gemma Nemotron
(104), .
; $(13, 13, 11, 1, 22)$ 0. Qwen
; . ( )
— Qwen 74 Gemma Nemotron 99 , —

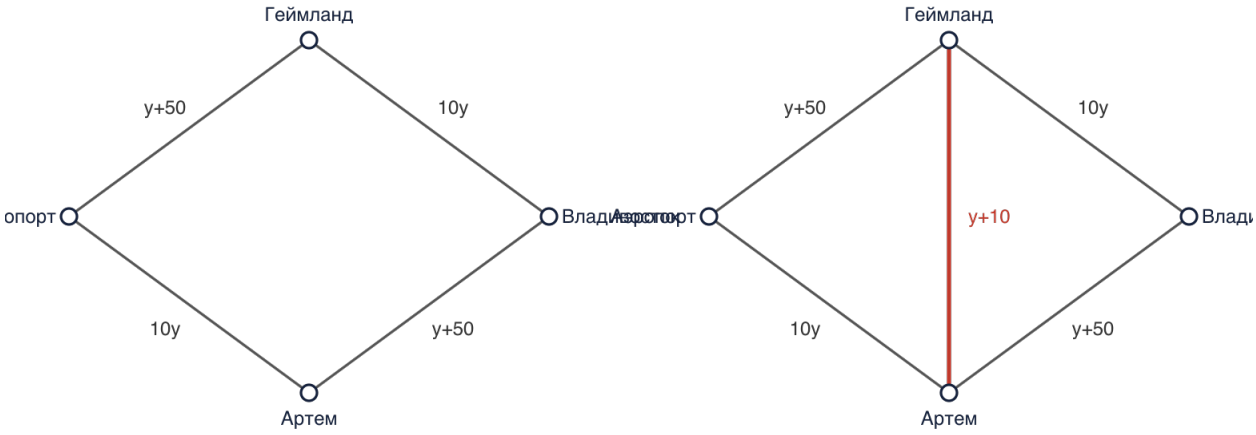

Figure 2: . ( ) – , 83. ( ) (, $y + 10$) « » – – – ; 2/2/2 92 — , .

Qwen (5–7) (1–3), , .

: **LLM-** , , , . — ; , ( ), ( ) ( ), — , , , — ($p = 2/3$) (;), . , : « » , - , , — totally mixed - [11]. (, ; ,), , « » . : , — , [1, 3], - - , [5].

— , . , , (100%- ,) «LLM- », , . : - Qwen Gemma Qwen — , . [7] , , [4]. , . / , ; ( ) ;

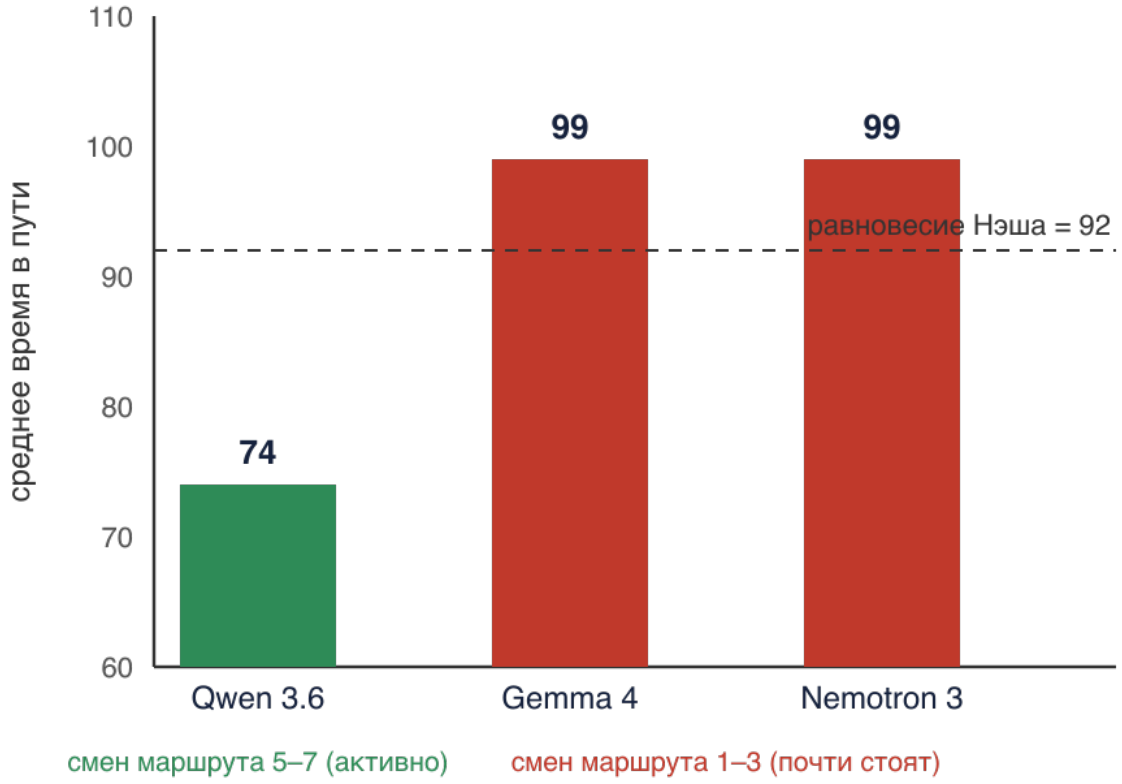

Figure 3: . (= 92,). Qwen (, 74), Gemma Nemotron (, 99). Qwen (5–7) (1–3).

; , human-AI [10], **human+AI**
. , .

: , ; ,
,
, , .
[2], , - .
; .
— , , , —
.

, , — ; — . ,
argmin , , , —
, , , . —
: , ,
, .

AICumene. V100 Thor;
.

References

1. Calvano2020
2. Fish2024
3. NoThreats2025
4. Einwiller2025
5. EquilibriumResiduals2026
6. Lu2024
7. InstitutionalAI2026
8. ScyllaExMachina2024
9. KLevel2024
10. DeliberateLab2025
11. Menshikov2017Markov