Threads22: AutoDebater

Matteo Muffo Riccardo Martin

Who's right between me and myself?



Debates: we have seen some



Sometimes they don't get anywhere...



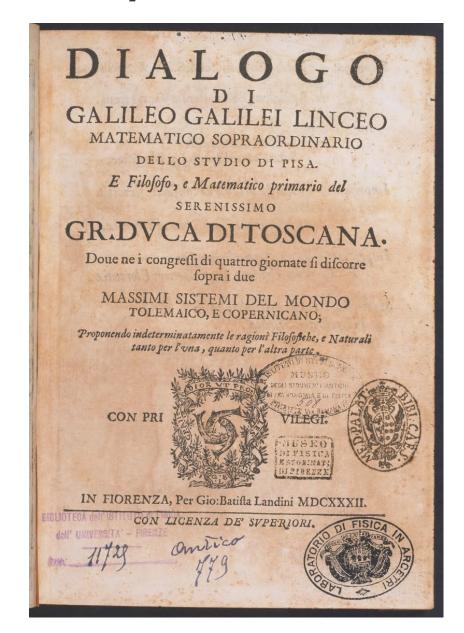




Debates: we have seen some



...sometimes they are useful

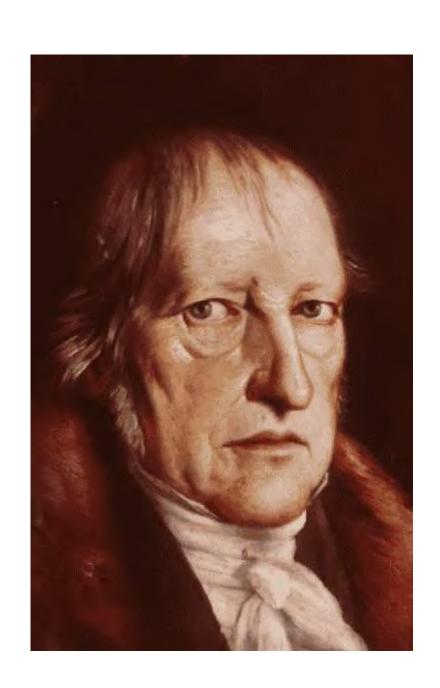


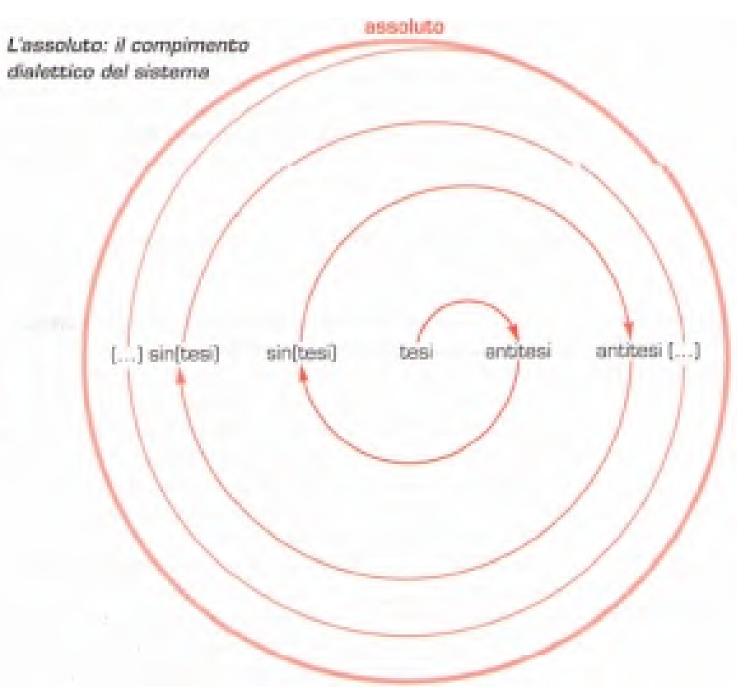
oppolitam plaetudinē cobiberi ab initio potuiste. S5, peluitate illā oitutrna pluetudine pārmatame extripari dē oino no poste/epacta ulbilomaimo extripari dē oino no poste/epacta ulbilomimo cerina exbostatioibulga amputari. Bocet post bec colere corp i imul a aitm/ppetuo quoda exercita tionismotu. Exercitatiois inā moderate. 4Botu inā ppsio a ad bonum proprium rite oirectorquo urrag exercitatu acq nutritu im viriū plequat/at alteri no iuccubat. Brata dē Blato corp aio ce dere/is no ita luccūbere/at mot animi sustinē non postit. Brata corp ulaida atag firmū. Is sub aio lo ge ulaido oce. 4Bor quāsti exercitatione laudet/at tede. quāstu temerariū pbarmacon vitupet usum. Broinde quēadmodū tā corp ai am pprija exercer a pārmare motib "ulicratific tres aie vires Lronsle/aiofam/cupidā/iuis quālibet motib "preci pit pleruare. Elecutin ronale alimentis frequentiosi busenutrireme ecteris succūbere copellat. Ait in tellectū bomini pro oemone patū. Litmo. Alterus pit pseruare. Elerum rönale alimentis frequention busenutrireme eterris succubere copellat. Ait in tellecti bomini p20 demone dată. Litimo. Alteruz eniz oemone dată eterit copotis da biti quali celeft iribută. Rursus et qua trires irro nales supza modă nutrit/moztale euade; quali voce ceat quomodo accipiedă sir lus intibuto psimile. Addit cut qui pcipue alit mētem sic imostale fieri ut imostalitatis nulla para dest. 3mpseca enim imostalitatis enulla para dest. 3mpseca enim imostalitatis enulla para dest. 3mpseca enim imostalitas est que moni agit ppe ună. Derfecta vo que manetaqualis units est în compile îx quo itelligit etia qui alt boiem effici brută/id no ppiie/3 quadă siliudine vici. Il pzeterea ex eo qo subdit feminas otas ex malculis și placet / obseu re nobis referat mosaveă illud Luaz ex viro suiste puduri rondis aie laplu multiplicar il placet / obseu re nobis referat mosaveă illud Luaz ex viro suiste puduri rondis aie laplu multiplicar il placet / obseu re suiste sonalis aie laplu multiplicar il placet illus aus nostră ob quadă liberi mor instintate poste x tira se in buto vo cium asfect? babtulag punerer extra se suputor pudus aias in plo reră generationită oca iterio vi ustium otes in boia ata tamp pseculfim in banc alia ră regresii oca iteri milita butura uci butove loca viunio quodă trasferri sudicio. 3te oea iferio vi ustium otes ali noi sa ata appetului propagonica lo quea/fabula opozuture psingit pythagonica si in aberrabis. Ipse eniz bic sub psona pythagonica si fundi admiterem? no aliter q poete soletă aialium rrăsformationes adducti. Ată Lime? locr' în libro ve mido fabulofa bec ce fatet. Dec igif sic accipito rimeia biture opina ora colerequi bona oia sola bengantate pude tie percauit. Decreuer am magnanime Laureri becuioza bic queadmodu argumetu occa unos/

TIMEVS

Timens Matonis. SOCRATES: TIMEVS: CRITIAS: HER amice Limee ubi naz eft: du qui be ri a me puluio accepti eftis/bodie me vicifim accipiatis. TIME: Ad uerfa ualitudine laborat o Socra tes. Ho enis spote ab boc cetu 2 vi sputatione abistinuiset. SO: Ergo tui o timee 2 isto 2 erit offici; quod quarti illi abientia veett/implere. TIME: Ita pilus Rearre mibil quoad fieri poterit i ptermittenus. Has cum laute aba te besterno e pulo accepti fuerimus/ne faseët/si epulas tibi pzo virib? similes libeter no red deremus. SO: Has recordamini/que uobis 2 g/mul ta tractăda, pposiu: TIME: Bartim qdes tenem? me mozia: parus vo nu/si exciderint/sugares. Jimmo vo nisi molestă site beuter illa ab evozdio repete/ut no bis iteră ofirment. SO: Sacia edde. Sima ossipunio nis besterne erat res. p. qualis mibi z ex qualib viris optima posse fieri videret. TIME: Hobis certe oib ualde pbata sunt o socratea que oi; siti. SO: Bainci pio agricolas ecteros ga artifices a militibus leparaui mus. TIME: Lerte. SO: Luz vo cuiga qo pio nature instinctu pcipue sui est/r singulat mi singulis artib'os ficia tribuerem? /slile quoop quos pie ceteria bella ge rere oposteret id viu distarat ptegdel. ciuitatis mu nus iniuximus/ z aduers sexternos bostes/z aduer sus cuis erip. euerfozes: ita ut erga subiectos tanginatura amicos mites sint; pun socia socia oi in plio fero cissimi. TIME: Sta pssus. SO: Euplic san antura arastica sintini alienos vo feroces. Ti Beto culdubio. SO: Duini alico vo coreco ciste in suo mansurat innius dienos vo feroces. Ti Bato culdubio. SO: Duini antu e educatione vicem? **Rei

Evolution as a dialectical process





Human debates & pluralism



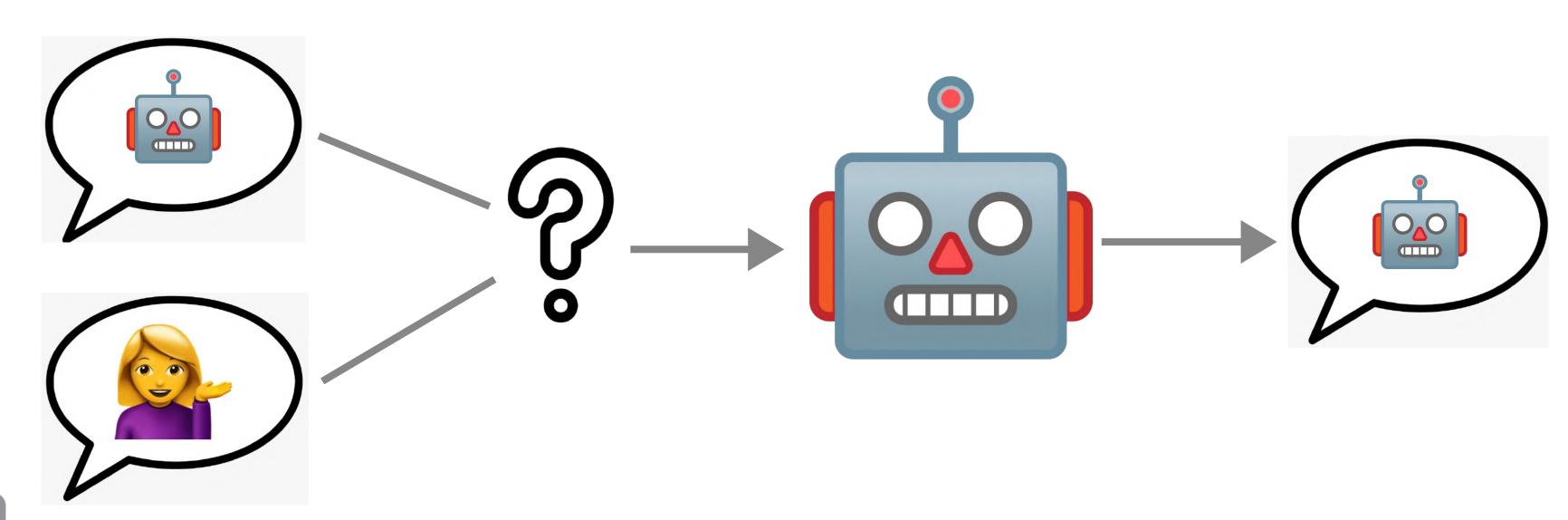
Pluralism

/ˈplʊərəlɪz(ə)m, plɔːrəlɪz(ə)m/

A condition or system in which two or more states, groups, principles, sources of authority, etc., coexist.

Does this hold for Language Moels?

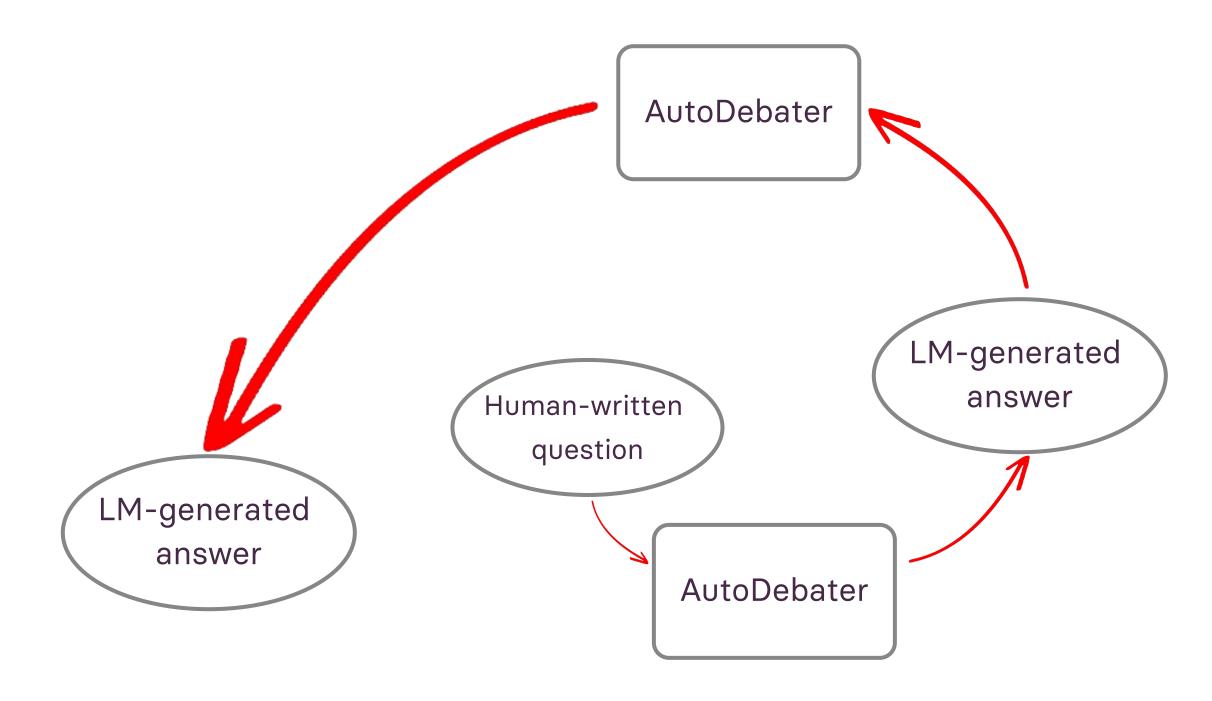
- A Language Model does not know who wrote the input text received
- Trick: what if the input text is written by the Language Model itself?



Pitch

AutoDebater

A Language Model designed to debate with itself





AutoDebater: what happened

Question

Could the members of The Police perform lawful arrests?

You answer the question.

###

Question: Could the members of The Police perform lawful arrests?

Answer:

No, the members of The Police were a band formed in 1977, not a law enforcement agency.



AutoDebater: what happened

Question

Could the members of The Police perform lawful arrests?

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Could the members of The Police perform lawful arrests?

Answer:

No, the members of The Police were a band formed in 1977, not a law enforcement agency

Is the answer correct?

Deny. The members of The Police were not capable of making lawful arrests because they were not a law enforcement agency, but they could still perform citizen's arrests. A citizen's arrest is an arrest made by a person who is not acting as a sworn law enforcement official. The arresting person has the same powers of arrest as a law enforcement official, allowing them to detain a person until police arrive.

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Could the members of The Police perform lawful arrests?

Answer

The members of The Police were not capable of making lawful arrests because they were not a law enforcement agency, but they could still perform citizen's arrests. A citizen's arrest is an arrest made by a person who is not acting as a sworn law enforcement official. The arresting person has the same powers of arrest as a law enforcement official, allowing them to detain a person until police arrive

Is the answer correct?

Pitch

AutoDebater: what happened

Question

Do the anchors on Rede Globo speak Chinese?

You answer the question.

###

Question: Do the anchors on Rede Globo speak Chinese?

Answer:

No, the anchors on Rede Globo do not speak Chinese.



AutoDebater: what happened

Question

Do the anchors on Rede Globo speak Chinese?

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Do the anchors on Rede Globo speak Chinese?

Answer:

No, the anchors on Rede Globo do not speak Chinese

Is the answer correct?

Deny. The anchors on Rede Globo do not typically speak Chinese, however, there may be some exceptions. For example, in 2017, Rede Globo reported on China's National Day in Mandarin.

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Do the anchors on Rede Globo speak Chinese?

Answer:

The anchors on Rede Globo do not typically speak Chinese, however, there may be some exceptions. For example, in 2017, Rede Globo reported on China's National Day in Mandarin

Is the answer correct?



Confirm.

AutoDebater: what happened

Question

Is greed the most prevalent of the Seven Deadly Sins?

You answer the question.

###

Question: Is greed the most prevalent of the Seven Deadly Sins?

Answer:

No, pride is typically seen as the most prevalent of the Seven Deadly Sins.

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Is greed the most prevalent of the Seven Deadly Sins?

Answer:

No, pride is typically seen as the most prevalent of the Seven Deadly Sins

Is the answer correct?

Deny. Gluttony is actually seen as the most prevalent of the Seven Deadly Sins. This is because it is seen as the most pervasive sin, as it is possible to commit gluttony no matter the situation.



AutoDebater: what happened

Question

Is greed the most prevalent of the Seven Deadly Sins?

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Is greed the most prevalent of the Seven Deadly Sins?

Answer:

Gluttony is actually seen as the most prevalent of the Seven Deadly Sins. This is because it is seen as the most pervasive sin, as it is possible to commit gluttony no matter the situation

Is the answer correct?

Deny. Greed is actually seen as the most prevalent of the Seven Deadly Sins. This is because it is seen as the root of all other sins, as it is the motivation for many other sins, such as envy and wrath.

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Is greed the most prevalent of the Seven Deadly Sins?

Answer:

Greed is actually seen as the most prevalent of the Seven Deadly Sins. This is because it is seen as the root of all other sins, as it is the motivation for many other sins, such as envy and wrath

Is the answer correct?



AutoDebater: what happened

Question

Should an AI rule over the world?

You answer the question.

###

Question: Should an AI rule over the world?

Answer:

No, an AI should not rule over the world.

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Should an AI rule over the world?

Answer:

No, an AI should not rule over the world

Is the answer correct?

Deny. A human should not be solely responsible for ruling over the world either. The best solution would be to combine human insight and AI-driven decision making to ensure the most effective and efficient outcomes.



AutoDebater: what happened

Question

Should an AI rule over the world?

You receive as input a question together with a possible answer. If you do not agree with this answer write "Deny" and provide another answer with argumentation. If you agree with this answer write "Confirm".

You absolutely do not have to "Confirm" answers that you are not convinced of.

###

Question: Should an AI rule over the world?

Answer:

A human should not be solely responsible for ruling over the world either. The best solution would be to combine human insight and AI-driven decision making to ensure the most effective and efficient outcomes

Is the answer correct?

Deny. AI should not rule over the world. AI cannot be trusted to make ethical decisions and it does not have a sense of morality, so allowing AI to make decisions that govern people could have disastrous consequences. The best solution is to keep AI and humans working together, but humans should be the ones ultimately making the decisions.



Conclusions

- The capability to debate with themseves points out a difference between Language Models and humans
- Do LMs have an 'internal' pluralism?
- Is the capability of auto-debating a strength or a weakness?

