

~~The goal:~~

~~Get Fit to make a completely fake video about fully automated AI swarms on 2b2t~~

~~Updated goal:~~

~~Get Fit to make MULTIPLE completely fake videos about fully automated AI swarms on~~

~~2b~~

~~LOL~~

Goal: completed

REQUIRED VIEWING:

https://www.youtube.com/watch?v=pFlc_jizgl0

<https://www.youtube.com/watch?v=LPCPUwL6MjQ>

<https://www.youtube.com/watch?v=No5SugTgleg>

Timeline (outdated, no longer relevant):

- ~~1. Contact fit about a potential follow up. Points to cover include:
 - a. Good job of the baritone video, it got a ton of views, one of your top 10 most viewed videos ever
 - b. Are you interested in a follow up?
 - c. Vercellino (who I mentioned before) has been working secretly on a modification to Baritone for some time now
 - d. Previously he's been making additions to the code, but he's gone closed source
 - e. I don't know what he's working on but I'm cautiously optimistic it will be amazing~~

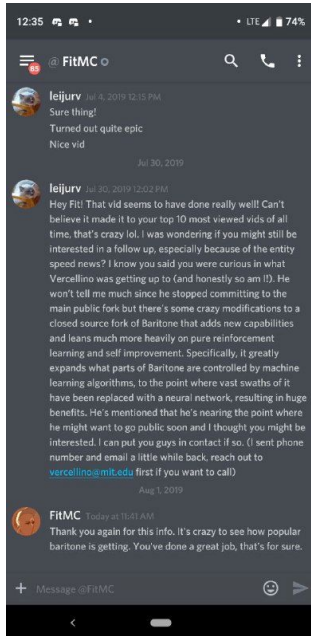
- ~~f. It relates to vastly expanding what parts of Baritone are controlled by machine learning algorithms, to the point where vast swaths of it have been replaced with a neural network~~
- ~~g. Vercellino can be contacted (here). (previously I gave fit conner's phone number and vercellino@mit.edu) Could give discord maybe?~~
- 2. Parkour ascend PR gets merged that evening
- 3. Fit doesn't buy it or care enough
- 4. Herm fakes a screenshot of someone asking him about this, sends on that chat
 - a. Perhaps at this point there could be footage of bots converging on someone, from the point of view of being killed by them?
- 5. Then asks me, and somehow I connect the dots to what conner is doing with parkour ascend maybe? Like I respond with "wait conner just added parkour ascend to baritone yesterday... is this a secret feature from his branch". The idea being that this is part of conner's plan to go public.
- 6. Fit contacts vercellino
 - a. Shares some details of what's going on with Fit
 - b. This is secret and goes against leijurv's vision for baritone. Please don't share this with him for now. I know I can't keep this a secret forever.
 - c. *Just an idea, I'm not sure if this would work out, but what if conner was like "i've been trying to get in contact with you for quite some time fit regarding this project without leijurv finding out about it, but I think I've been secretive enough that it won't matter how much he already knows. I know he won't like it, and if he knew exactly what I was doing he would never have put us it touch"*
 - d. It's a reinforcement learning network that's trained on a system of rewards and penalties
 - i. Penalty for dying, penalty for being disconnected
 - ii. No rewards
 - e. Introduce minerl, etc
 - f. All of this shit:
 - i.
- 7. Continued dialogue between fit and vercellino, maybe some footage is provided
 - a. Show off multiple bots killing targets
- 8. Then the real test, to know if this should continue. **I will ask fit what vercellino is telling him. If fit tells me the truth, we give up on the idea. If fit tells me a lie that vercellino feeds him, we're in business**
 - a. Obviously not until they've talked for some time and established a rapport etc

^ haha yes this is actually working fit said he'd tell me everything conner said but he has said NOTHING

For-realsies initial message to Fit, as sent (timeline sequence #1)

Hey Fit! That vid seems to have done really well! Can't believe it made it to your top 10 most viewed vids of all time, that's crazy lol. I was wondering if you might still be interested in a follow up, especially because of the entity speed news? I know you said you were curious in what Vercellino was getting up to

(and honestly so am I!). He won't tell me much since he stopped committing to the main public fork but there's some crazy modifications to a closed source fork of Baritone that adds new capabilities and leans much more heavily on pure reinforcement learning and self improvement. Specifically, it greatly expands what parts of Baritone are controlled by machine learning algorithms, to the point where vast swaths of it have been replaced with a neural network, resulting in huge benefits. He's mentioned that he's nearing the point where he might want to go public soon and I thought you might be interested. I can put you guys in contact if so. (I sent phone number and email a little while back, reach out to vercellino@mit.edu first if you want to call)



This response basically indicates disinterest, we need some kind of way to make him take it seriously.

Some ideas include making fit feel like he is discovering this instead of me telling him about it. For example, start vague rumors from conner of changing pvp on 2b for ever, but with no specifics, not even mentioning it's based on baritone.

Perhaps this "release announcement" or whatever from conner could be brought up to fit with people asking fit if he knows anything about it.

A long time scale is key on that.

The option of last resort is me "burning" this communication channel, by disavowing conner's actions.

This would be saying something like:

<cut/pasted somewhere else>

Another fit response out of the blue on Aug 8

The A.I. video just surpassed 1 million views. I bet baritone interest is at an all time high at this point

This is pretty much certainly expressing interest in baritone and potentially a follow up. Why else would he send this?

A potential response could be

The question is if I should also mention conner's info....

"they know crystal pvp, and they really don't look like baritone, see how they do sprint jumps at an angle up blocks? baritone doesn't do that"

Realized that fit doesn't know that conner=c0nn3r=vercellino

Sent to FitMC:

yeah it's crazy!

baritone has been downloaded 178,416 times from 12:01am pst july 4 to 12:01am pst aug 8, and still getting consistently 4 to 5 thousand downloads a day. i don't know what the saturation is among 2b players specifically, for example only 92,310 of those downloads were for 1.12.2. society did a survey some months back and only 20% of players reported using baritone, i would estimate it's probably higher now, maybe around half

if you're interested in a follow up i highly recommend speaking to conner vercellino, i gave you his email previously but he just shared his discord with me in case that's easier it's @c0nn3r#7254, he just contributed something to baritone's codebase from his private branch and it's pretty cool, it's actually the first new movement in the last year that baritone has been taught how to do

<https://youtu.be/9Fb5tdsIIJc>

<https://youtu.be/bOyS-3FVHIw>

Faked screenshots sent, gambino on board, etc

At this point in time, I need to give an update to both herm and fit, as I confirm my suspicion (that i expressed to herm, NOT to fit!) that these bot swarms are vercellino's doing. Therefore my message to herm needs to be fairly simple, like "I confirm it and it's true" is all it really needs to say, but the message to fit needs to introduce the whole thing, like:

To herm:

conner hasn't responded. did you ask if any of the bots were elon_musk or septave? that would be a dead giveaway that conner's behind it if so, I could go to him with that
(sent)

<https://www.dropbox.com/s/bhp344aozj5zzhi/botpov.mp4?dl=0>
<https://www.dropbox.com/s/71b01qaqen6i0v3/victimpov.mp4?dl=0>

Sent uwu

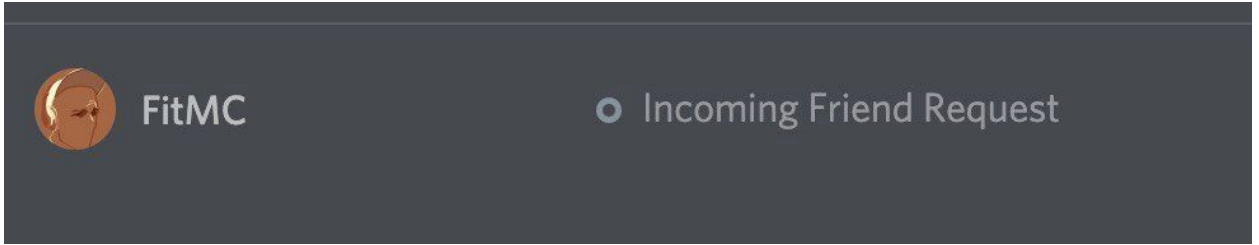
(via streamable -> constantiam chat -> whatever)

Going to help them make the connection of conner <-> vercellino by sending a telegram screenshot including the full name

DM confronting conner:

- someone just sent me this <https://streamable.com/qa1xq>
so this has been fucking you all along
if you actually wanted to hide it you shouldn't have used elon_musk i remember it from the sophomore year minecraft server
i literally do not understand why you have been ignoring me, did you think i would never find out
and i suppose this why you added that parkour ascend movement then?
just to improve your bot swarm in bedrock, just like this?
god
- I knew it was only a matter of time until someone posted a video.
- I'm surprised it took this long.
- Now that it's out there I'm not worried about hiding, I'm targeting to fill a significant portion of the slots on 2b2t with my army.
- You should have joined me when you had the chance.
- I'm going to make Baritone look like a toy.
- And yes, I'm very proud of the bedrock jump improvements, the network caught on quickly to that new training data and surpassed baritone handily...
- The only thing left to do let the bots train, and 2b2t is the only environment that can provide something hard enough for these bots, densely populated with other intelligent yet adversarial actors with which it could communicate. I hope these bots can learn diplomacy, and the value of peace but I don't know if they will. I'm using craftassist's natural language processing to give it a voice that I've hooked up through /msg but it hasn't used it yet. But I shouldn't tell you too much... fit didn't contact me about your silly

baritone spawn escape a month ago, but look at this:



- The swarm is clearly superior. I won't be sharing any details with you, the only thing that could ruin this at this point is someone stealing my research under me. But if fit continues to put out videos on this - I will have won. I'm going to do whatever needs to be done to complete my dissertation.

Research thing is lame af
IDK what else to do, but research is
much on cloud servers i can't stop now i need to finish my research

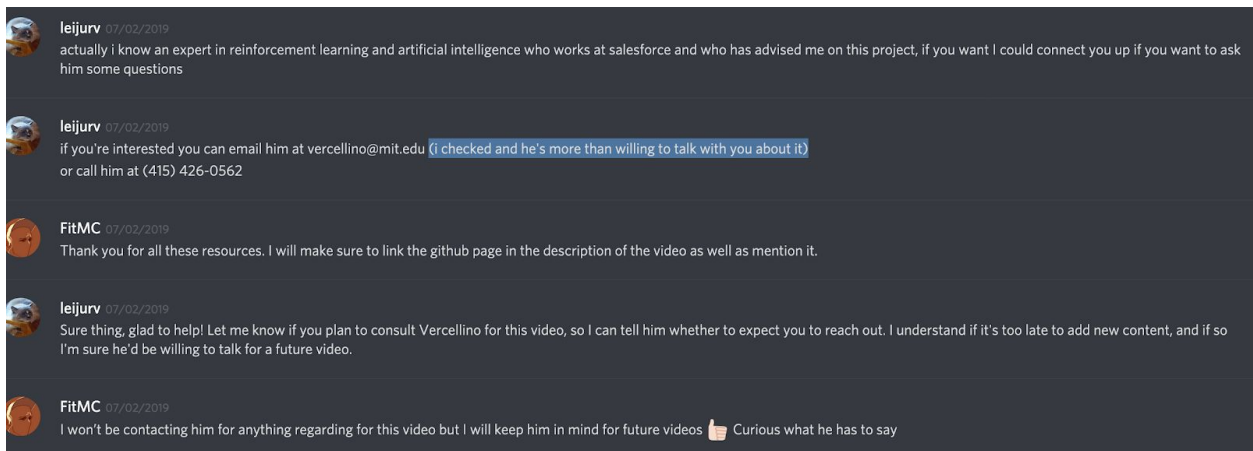
Also should bring up how I asked you if you'd be willing to talk to fit in early july and you said yes.

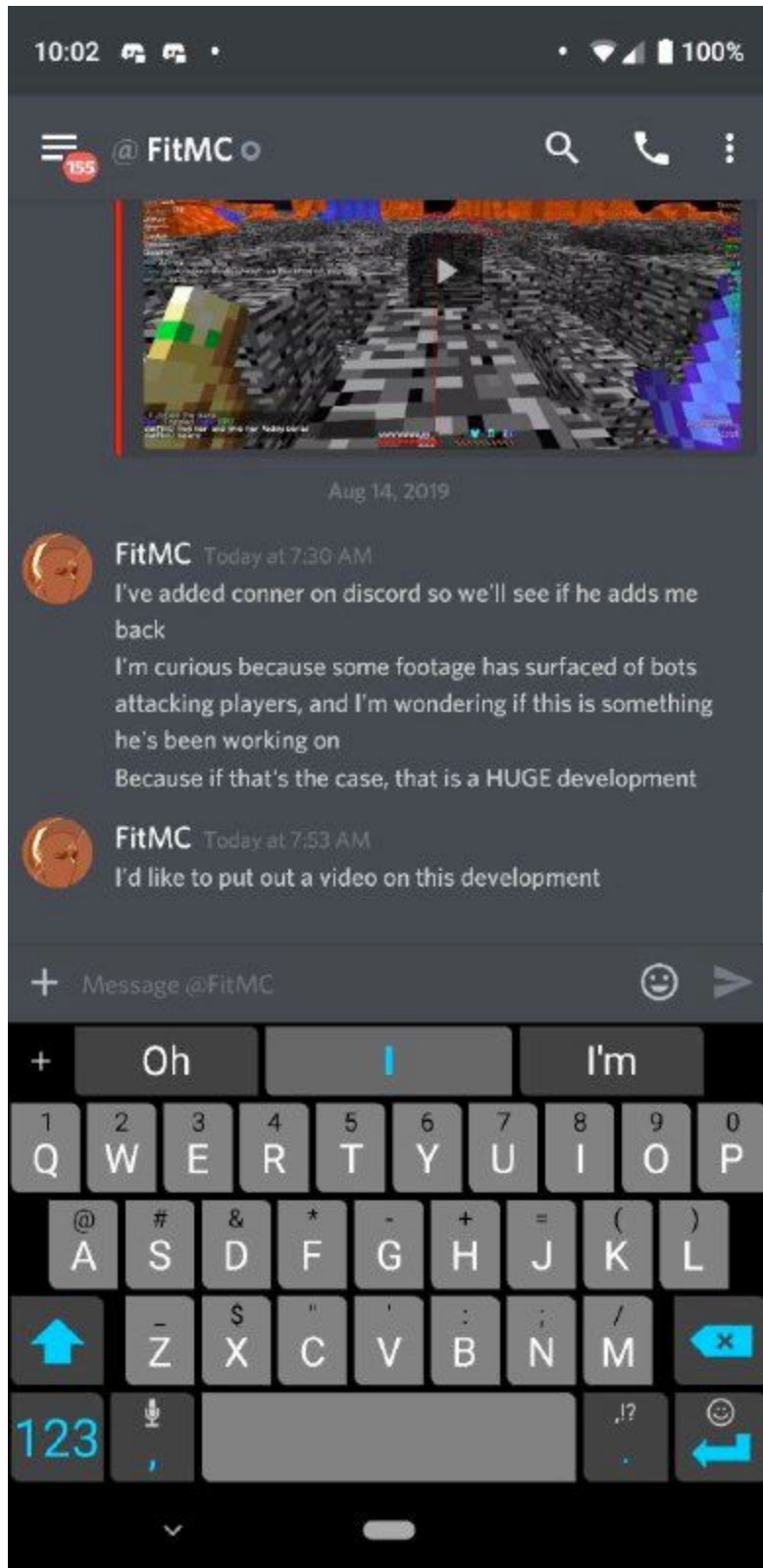
Something like **“fit didn't contact me about your silly baritone spawn escape a month ago, but look at this *screenshot of friend request from fit*. The swarm is clearly superior”**

Oh OK, yes I like it

ohh are we going to include the neural network visualization? seems

Remember that in the narrative you already know fit is interested:





<https://www.youtube.com/watch?v=LPCPUwL6MjQ>

So he fucking posts it

Here's the question: did he really expect conner to respond in just 4 hours? Or did he anticipate it would be too soon, with the thought of making a follow up video?

SENT to fit:

i see you made the video. never mind on this. i have learned more about the project and i do NOT support this. i think you saw what i sent in spawnmason #general? vercellino is creating these bot swarms, and using reinforcement learning to teach the bots combat autonomously. he isn't basing this on baritone or any existing combat techniques, this is learning from scratch using self-play and self-improvement based on minerl (<http://minerl.io/>), look into it. they coordinate with each other under the control of a central custom client, and i have no idea how he managed this. just now he bragged to me that he's also using craftassist (<https://ai.facebook.com/blog/craftassist-platform-for-collaborative-minecraft-bots/> <https://research.fb.com/publications/craftassist-a-framework-for-dialogue-enabled-interactive-agents/>). this is horrifying, and I don't want it to become the new reality. i'm not sure how much you know or who told you (herm told me), but you should not promote this. did you talk with conner before this video? if you promote it, he gets more visibility and potentially more funding for this research and his servers. at the same time, this will probably get out anyway which is what kills me. any technical details he gives you, could you share with me? need to counter or even reconstruct this, i believe that might be the only thing to destroy his motivation to pursue this perhaps

FitMCToday at 2:47 PM

I haven't spoken to conner at all over any of this. I assumed you two were friends. If he speaks to me, I can let you know what he says.

I thought he was helping you with baritone but it's clear he has his own intentions

FitMCToday at 2:56 PM

But I will help you any way that I can

i do know him in real life yes. he has not shared anything but the barest details of this with me, and i only learned yesterday what he was truly doing with his research. while he has helped with baritone, for example in the contribution of the parkour ascend movement, and earlier in development, i have now realized that pretty much all of this was a means to an end of creating his bot swarm, not having the project's best interests at heart at all. he's basically betrayed the project and what we are working towards with baritone (not to mention him always starting drama in baritone's issues). he's interested in money and fame and making a name for himself in the ml community, while baritone is interested in free and open source software that's accessible to all. i've removed his baritone rank and he no longer has access to the

development chat. i assumed you had reached out to him though, you just said "I've added conner on discord so we'll see if he adds me back" did he?

https://media.discordapp.net/attachments/587173134965276672/611317986128560147/Screenshot_20190814-123720.png

https://media.discordapp.net/attachments/587173134965276672/611317996513656832/Screenshot_20190814-123727.png

https://media.discordapp.net/attachments/587173134965276672/611318007335092226/Screenshot_20190814-123733.png

https://media.discordapp.net/attachments/587173134965276672/611318016252313647/Screenshot_20190814-123738.png

thanks for offering to help

FitMCToday at 3:03 PM

This is like Dr. Evil level stuff, sheesh

He accepted my friend request but we haven't spoken yet

leijurvToday at 3:19 PM

ugh... could you continue as you were, with gathering information relating to making a video? maybe talk to conner and try and get information from him, how was this developed, is anyone else helping, does anyone else have access, how does it actually work, what was it trained on (data and hardware), what are his plans, will he ever release it, etc

FitMCToday at 3:26 PM

Will do, I will tell you everything

Response from conner:

Hey Fit. Just watched the vid, nice. Impressive how you figured out it was me so fast. Those bot tests were indeed my doing. I've been working on a research project for some months now built around bots studying and learning how to play advanced, high level minecraft. The clip you showcased was one of my combat test nets.

I have an idea on how we can benefit each other greatly. The best possible thing for my research is attention, and you as a youtuber are after the same thing.

So I'll cut to the chase: since someone dropped footage I didn't expect, and you jumped on it faster than I could've imagined, I'll give you the rest: Footage of the attack on user xdtfwew from the bot's training POV, and more early peeks into the training for my PVP AI fleet. In exchange, it should be made clear that ALL of this work has been authored by me and me alone, and that this is the future of Minecraft PVP; the next wave. Most importantly of all, I want you to do it justice, so to speak. Don't gloss over how it works by just saying that it is "machine learning", and don't rush to put out a video that's missing most of the information. I understand, of course, that you can't just make a technical showcase video. I will explain this in a way that benefits everyone. If this becomes publicity for me, I don't want a dumbed down "the bot learned from baritone". I want you to actually explain (at a high level, an understandable level of course) how it works.

I'm sure your audience would of course also want my backstory, motivation, history with leijurv, how we got to this point. But I want to make it clear that while I'm willing to share that with you, it is not my primary goal.

^ edited, see the real message down below

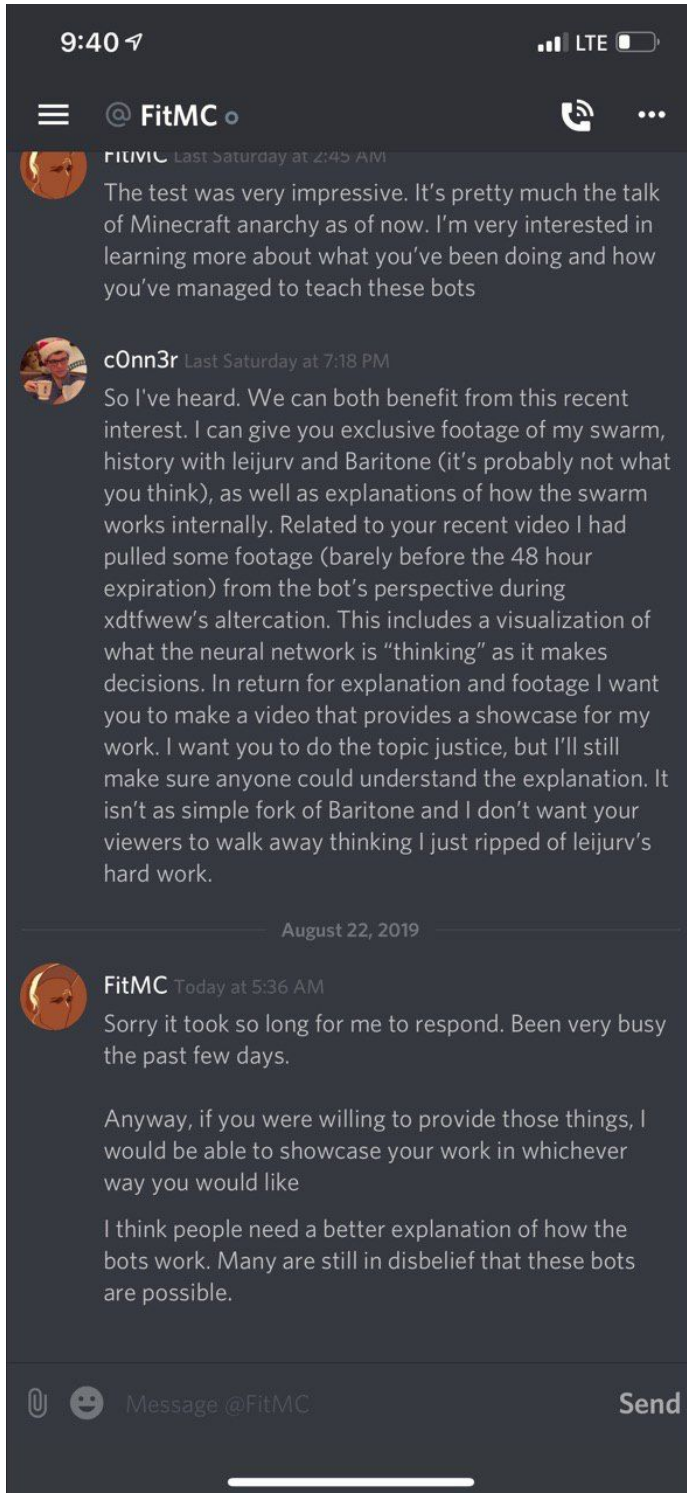
FitMCToday at 2:45 AM

The test was very impressive. It's pretty much the talk of Minecraft anarchy as of now. I'm very interested in learning more about what you've been doing and how you've managed to teach these bots

HERE TK

We can both benefit from this recent interest. I can give you exclusive footage of my swarm, history with leijurv and Baritone (it's probably not what you think), and explanations of how the swarm works internally. Related to your recent video I had pulled some footage (barely before the 48 hour expiration) from the bot's perspective during xdtfwew's altercation. This includes a visualization of what the neural network is "thinking" as it makes decisions. In return for explanation and footage I want you to make a video that provides a showcase for my work. I want you to do the topic justice, but I'll still make sure anyone could understand the explanation. It isn't a simple fork of Baritone and I don't want your viewers to walk away thinking I just ripped of leijurv's hard work.

Come on, the dating thing would be so fucking funny



^ I LITERALLY CANNOT FATHOM A BETTER RESPONSE FROM FIT ^

He is willing to showcase the work, he agrees people need an explanation, he realizes that people don't believe it's possible but he still does. POG.

Alright its FINALLY time to dump all this info!!

Here begins the big boi essay

No worries about the slow reply, totally understand. Fantastic, glad to have you on board.

Right off the bat I think I'd like to clarify the biggest misconception that people got from your last video, which is the cause of all this skepticism on the subreddit. My bots are not a simple modification of Baritone with some new features — my "bots" are a singular neural network whose training set includes recorded actions that Baritone took; while running there is no Baritone code. It would be like saying that a neural network trained to identify images is actually human, when really all that happened was a human labeled its training set.

I added parkour ascend to Baritone to address a gap in the network's training set. The network learned from the new data, that it's able to jump and ascend over gaps, and refined on its own without Baritone. All it needed was an example. The MineRL dataset has no bedrock level pathing, just standard Minecraft (non-anarchy) gameplay. I therefore augmented this dataset with recorded Baritone actions. But I'm getting ahead of myself, let's talk big picture.

With the entity speed patch, the community has realized that player automation is now more useful than ever. See


https://www.reddit.com/r/2b2t/comments/cg9fh1/update_on_whats_worked_on_and_whats_incoming/eugf47m/

https://www.reddit.com/r/2b2t/comments/cgmfhu/benefits_of_patching_entity_speed/euj4dlk/

https://www.reddit.com/r/2b2t/comments/cgrl35/entity_speed_gone_better_travel_ai/

https://www.reddit.com/r/2b2t/comments/cggl1n/entity_speed_patch_will_launch_a_2b2t_dark_age/

Now, I think leijurv should capitalize on this huge new demand, and add monetization. He doesn't want to. He's too preoccupied with the short term benefit of the community that he doesn't realize the long term; where this is all eventually heading regardless of what he or anyone else wants.

 c0nn3r commented on Jun 1 Contributor + 😊 ...

Follow up to a conversation with @leijurv and @nacgarg on Telegram. This is meant to summarize my general ideas and provide a few action items for next steps for our goals.

Sparking this issue were two events: Baritone has been trailing FAR behind Future (@0-x-2-2 🙄) in recurring revenue (Baritone has only received a few one time donations) and @leijurv has disappointingly taken an internship instead of working on Baritone full-time during the summer. This is despite Baritone being a far more powerful tool than Impact or Future. @ZeroMemes has also opted not to share Impact donations with @leijurv. The goal of this issue is to lay the groundwork for a sustainable revenue model for Baritone - allowing full-time development for at least one developer. Potential ideas discussed are laid out below:

🔥 Baritone Premium™

We could gate a few Baritone features behind a recurrent membership fee and introduce a so called "Baritone Premium". A few ideas for what features could be moved to the premium version are below:

- Parkour Functionality
- Path Color Customization
- Custom Cape
- Builder (more on this later)
- Auto farmer
- Freecam Pathing
- Tunneling / Clearing Areas

🏃 Charging per a Block Traversed

Users would have a limited number of blocks they could traverse using the path finding functionality per a month before they would be asked to pay a few cents for each 100 blocks walked.

⚡ Surge Pricing

Building off of the idea to charge per a path, we also discussed dynamically pricing different paths (2B2T is almost exclusively our target market here). If you wanted to move along the nether highway for instance - that would cost more based on how popular that destination was among the player base. Pricing would also be more expensive based on the time of day that path was popular during.

💰 Market-based Builder

Baritone's builder functionality is about to receive a few much needed upgrades (being able to use chests as a source of items, in-game previews to name a few), one option is just gating these behind premium - another is having players pay per a build. Each block placed would be priced based on its blocks popularity within Baritone built structures. This would cause items such as wool blocks to have a higher price than say cactus. Besides promoting diversity within player creations, this would also provide a marketplace where players could select from a library of previous structures that they could then pay to build in game.

👉 Next steps

@nacgarg has offered to write the backend and @c0nn3r offered to do both the web frontend and surge pricing strategy.

We look forward to improving Baritone's sustainable development. 🙌

@c0nn3r and @nacgarg

👍 8 🗨️ 25 😊 3 🏃 4 😊 6 🚀 7

Leijurv has been continuously developing baritone for the last year. If you look at that "bots" video on his channel that you pulled some footage from, you can see that he had control of multiple bots working together all the way back in November of last year... but he immediately scrapped it. No videos on it have been released since then, the feature has not been present in any releases. It's not that he gave up or lost interest: he's been actively developing other parts of baritone, daily! And this is clearly a killer feature, being able to simultaneously control many accounts.

He killed off the feature because he's scared of its potential. But I'm not. While he recognizes the video is cool and he shows it off, he hasn't released the source code and I know

for a fact he hasn't touched it since. Personally, I was blown away by that video and the huge potential of something like that. It's been very difficult to reconstruct from the threads he left behind in the codebase, but I've created something that goes above and beyond that clip. While he claims to love open source software and we've argued intensely about what to do with Baritone, he never released this bot-system branch.

Now to address the "disbelief" and explain how this actually works. It's not magic. I've seen some posts ranging from people recognizing how plausible it is to create video game playing bots using machine learning:

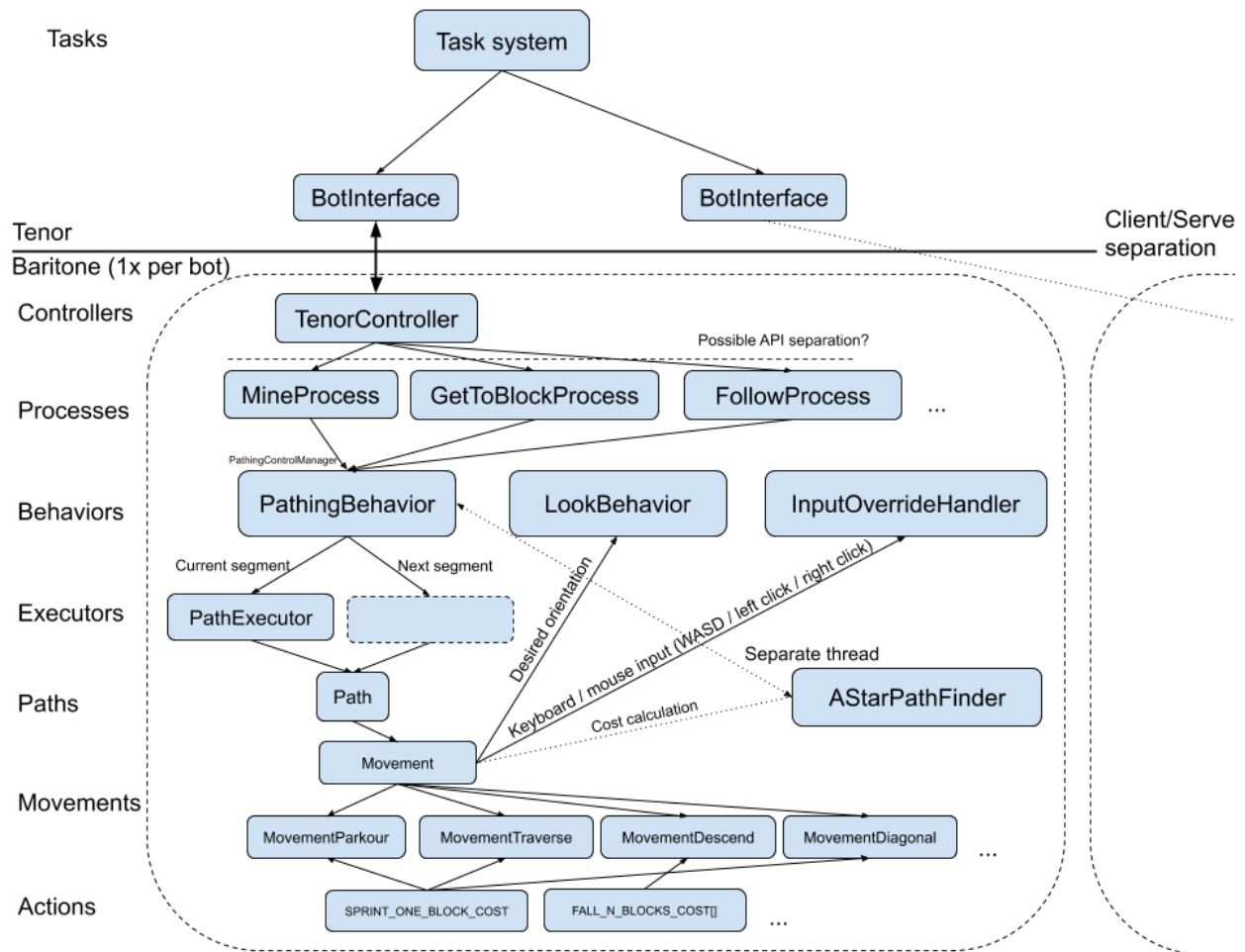
https://www.reddit.com/r/2b2t/comments/ctof1m/ok_so_basically/

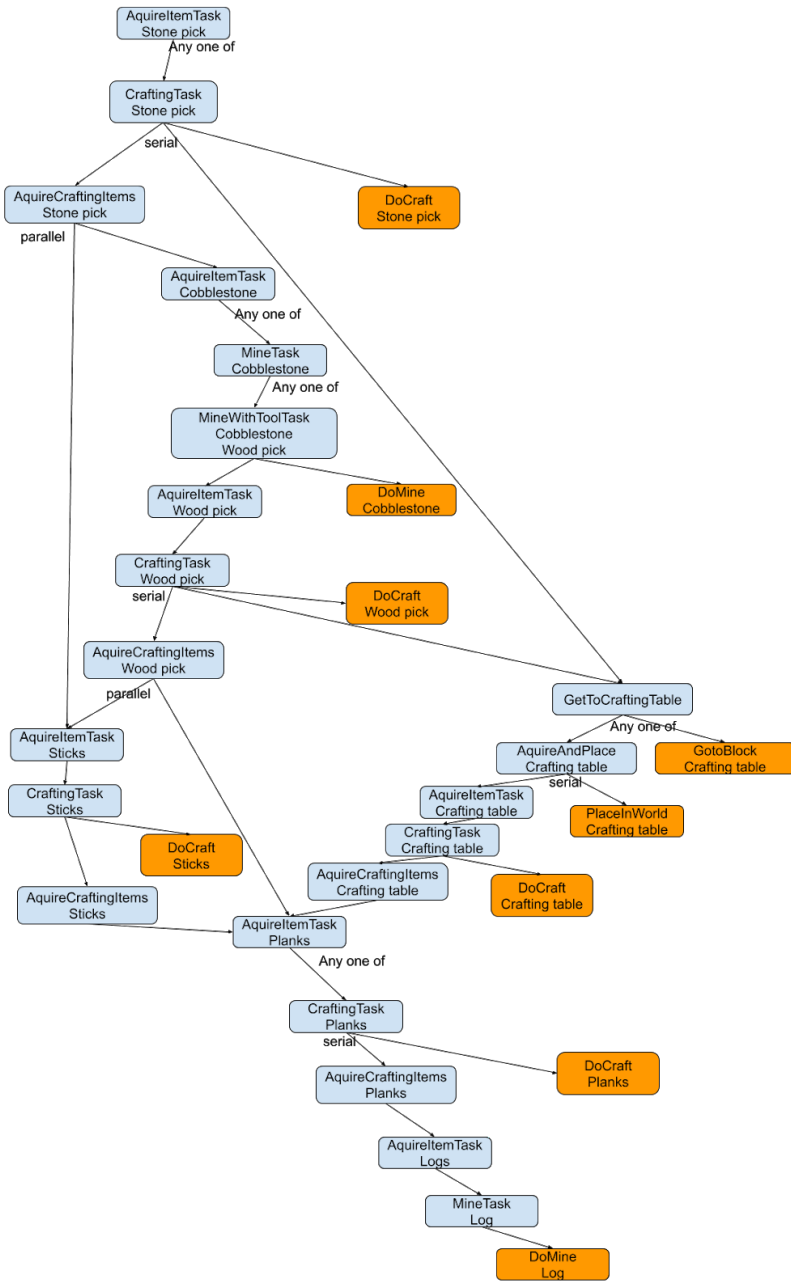
To some others (correctly!) pointing out that Baritone is inadequate to be truly competitive in PvP:

https://www.reddit.com/r/2b2t/comments/ctcvy9/shitty_shitpost/

But, this is not Baritone.

Baritone is a robot with unadaptable defined structure and code (see image of Baritone's internal layout). My bots are an elastic neural structure, with no fixed layout.





Alright, I guess I should ask how familiar are you with reinforcement learning and where it's been applied? There have been some high profile applications of it to games in the last few years, such as <https://deepmind.com/blog/alphago-zero-learning-scratch/> https://en.wikipedia.org/wiki/AlphaGo_Zero and <https://chess24.com/en/read/news/a-new-age-in-computer-chess-leela-beats-stockfish> and <https://openai.com/five/> <https://www.theverge.com/2019/4/13/18309459/openai-five-dota-2-finals-ai-bot-competition-og-e-sports-the-international-champion> https://en.wikipedia.org/wiki/OpenAI_Five

So, anyone who thinks this is categorically impossible is quite behind the times.

And there are also some applications to Minecraft. Minecraft has had it harder due to how complicated the game is. This is commonly called a sparse reward hierarchy. There is no way to get the “reward” of progressing to mining iron, unless you’ve already got a stone pick (this is what makes it a hierarchy). And it’s very sparse. You have to be at the correct position in the world to collect certain rewards (there’s no way to punch a tree in a desert). The main challenge is Minecraft’s 3-D exploration, and lack of dense information (you only know what’s in the direction you’re looking).

Nevertheless, there has been a surge of success in the last few months that I’m riding on top of. First, MineRL came out in April, the product of a team of 12 researchers at Carnegie Mellon, Microsoft Research, and the University of Maryland. Leijurv may have mentioned that I stopped committing my changes to the public branch of Baritone, it was at this time that I realized that what I dreamed off was truly possible. Look at <https://arxiv.org/pdf/1904.10079.pdf> and <http://minerl.io/competition/> and <http://minerl.io/>. V2 of their paper was released July 31. This paper (reinforcement learning to autonomously play Minecraft to the point of obtaining diamond) was accepted to be presented at https://en.wikipedia.org/wiki/Conference_on_Neural_Information_Processing_Systems Anyone who believes that it’s not possible for a neural network to control a Minecraft player just needs to look at this paper to realize that it’s been possible for months.

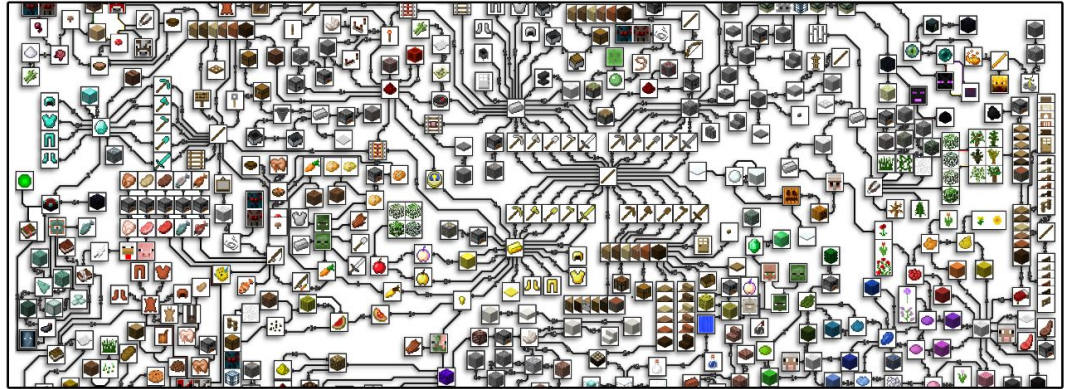


Figure 1: A subset of the Minecraft item hierarchy (totaling 371 unique items). Each node is a unique Minecraft item, block, or non-player character, and a directed edge between two nodes denotes that one is a prerequisite for another. Each item presents its own unique set of challenges, so coverage of the full hierarchy by one player takes several hundred hours.



Figure 2: Images of various stages of six of seven total environments.

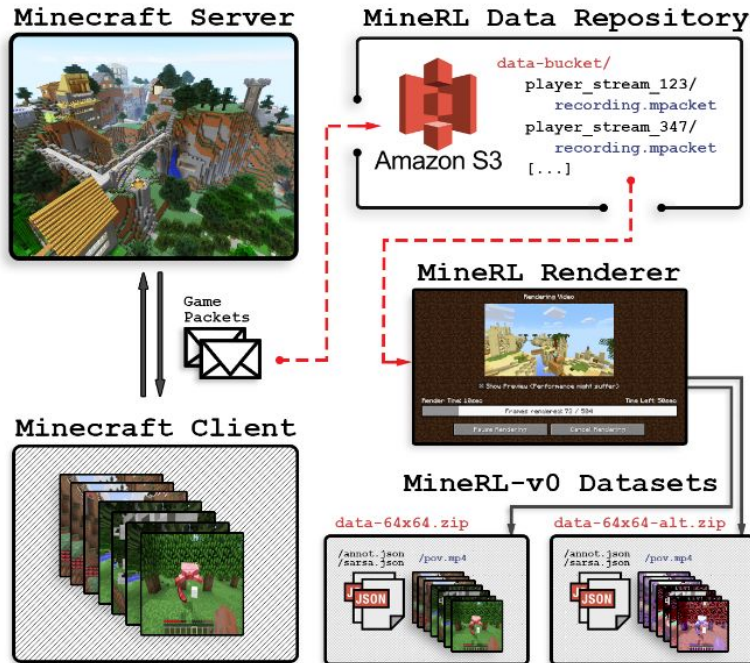


Figure 3: A diagram of the MineRL data collection platform. Our system renders demonstrations from packet-level data, so we can easily rerender our data with different parameters.

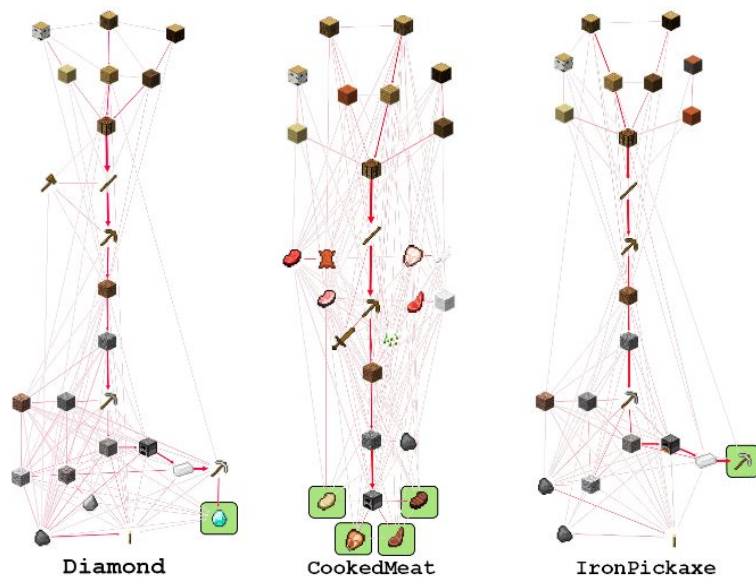
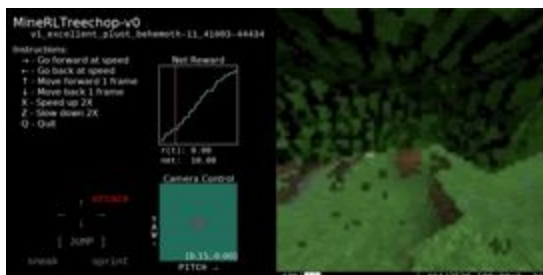


Figure 5: Item precedence frequency graphs for ObtainDiamond (left), ObtainCookedMeat (middle), and ObtainIronPickaxe (right). The thickness of each line indicates the number of times a player collected item A then subsequently item B .



Further work has also been done by facebook:

<https://research.facebook.com/wp-content/uploads/2019/07/Why-Build-an-Assistant-in-Minecraft-v3.pdf>

Taking inspiration from the MineRL project, I start by giving it footage of skilled Minecraft gameplay. In the first phase of training, which I call "learning by example", the only goal is to emulate recorded behavior. This is also commonly called "imitation learning".

As the MineRL paper puts it:

Many of the recent, most celebrated successes of artificial intelligence (AI), such as AlphaStar, AlphaGo, OpenAI Five, and their derivative systems, utilize deep reinforcement learning to achieve human or super-human level performance in sequential decision-making tasks. As established by Amodei and Hernandez [1], these improvements to the state-of-the-art have thus far required exponentially increasing computational power to achieve such performance. In part, this is due to an increase in the computation required per environment-sample; however, the most significant change is the number of environment-samples required for training. For example, DQN [13], A3C [14], and Rainbow DQN [9] have been applied to ATARI 2600 games [2] and require from 44 to over 200 million frames (200 to over 900 hours) to achieve human-level performance. On more complex domains: OpenAI Five utilizes 11,000+ years of Dota 2 gameplay [18], AlphaGoZero uses 4.9 million games of self-play in Go [23], and AlphaStar uses 200 years of Starcraft II gameplay [5]. Due to the growing computational requirements, a shrinking portion of the AI community has the resources to improve these systems and reproduce state-of-the-art results. Additionally, the application of many reinforcement learning techniques to real-world challenges, such as self-driving vehicles, is hindered by the raw number of required samples. In these real-world domains, policy roll-outs can be costly and simulators are not yet accurate enough to yield policies robust to real-world conditions.

One well-known way to reduce the environment sample-complexity of the aforementioned methods is to leverage human priors and demonstrations of the desired behavior.

So why have I had far more success than the original MineRL paper? Because I augmented their training set. They have publicized their set of over 60,000,000 individual “examples” (meaning “in *_this_* situation, while working towards *_this_* goal, a human player took *_this_* action (e.g. “pressed W while moving the camera to the left slightly”)). My key realization is that Baritone can generate this data, just by recording what Baritone does. And Baritone can be run much faster than a human. I generate massive amounts of training data from Baritone at over 200tps, with the game running at ~10x speed. It only renders one frame per in game tick to feed to the network. A human playing Minecraft at normal 1x speed is imperfect. Generating nearly perfect training data at 10x speed, without needing a human (so it can be run on many machines at once) is my key advantage over MineRL.

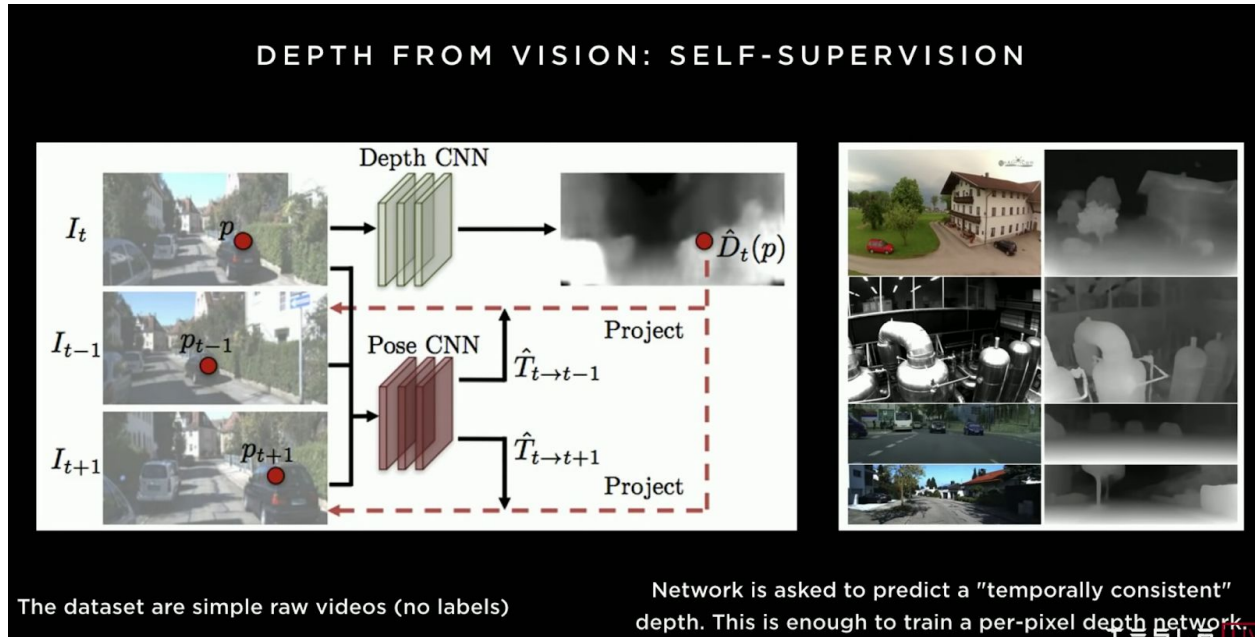
(As an aside: I suggested exactly this to William Guss, the lead of the MineRL project. He discounted the idea of using Baritone to generate training data, saying it wouldn’t work. William, you may remember me as elon_musk or septave, just know I was right.)

This is good enough for the simple tasks of MineRL. Each network they train has a certain goal, and control hands off to the next one once that goal is complete (e.g. their Treechop network is replaced by the ObtainWoodPickaxe network as soon as it’s done). I decided to create what’s called a “policy network” on top of the game control network. The policy network decides what high level goals to work for (“get to these coordinates”, “kill this mob”, “acquire wood”, etc). The crucial insight was that the policy network can be trained by pure unsupervised reinforcement learning, with the same inputs as the game control network. The key difference is that training the policy network doesn’t have a goal. I don’t tell it what to

do, it decides for itself what the game control network should be commanded to do, in order to maximize its reward and minimize its punishment.

For training this step, I took footage primarily from your channel, as well as some from other 2b2t youtubers. Not from your videos, but from your exploration streams. This works using a pose estimation network with a depth prediction network to generate the keypresses and mouse movements that were most likely carried out. See

<https://youtu.be/Ucp0TTmvqOE?t=8503> for the inspiration for this.



Keep in mind that this is only after transfer learning of the baritone pathfinding system to form the initial base for the network, so that it knows (in neural form) how to navigate minecraft (the game control network). The bot trains itself to generally replicate the actions you perform while exploring spawn, building, etc. This is a useful training step because it allows it to quickly and easily learn simple facts about Minecraft and the server specifically, without having to actually log on. (e.g. echests, how crystals work, shulkers, etc).

Then, it moves onto the second phase, which is where it finally actually takes control of a player in game, now that it's learned what to do. This stage is commonly called "refinement" (but I like to call it "perfection"). It plays through over and over, but now it has no example to learn against other than itself. Instead of trying to replicate the decisions of an imperfect human or imperfect pathfinder, it is trying to make whatever decisions are necessary to achieve its goal.

Figuring out what goal to give it was tricky. What do we actually want it to do, and what goal will help us achieve that? As an initial proof of concept, I decided to give it no rewards, only a punishment for death. It tries to minimize punishment over time, so when combined with the respawning mechanic this results in an AI that learns nothing but "survival" (more precisely, dying as infrequently as possible).

An issue that I ran into at first was that it would just decide to stand stationary at spawn. It realized almost immediately that moving in any direction in 2b2t spawn without a clear sense of what to do resulted into walking into lava and withers, or falling through a gap down to

bedrock (an example of what we call “dropout” in machine learning). I initially overcame this by adding a penalty for disconnection that is greater than that of death. This caused it to make the necessary movements to bypass the anti-AFK.

(I actually took a snapshot of the network as of this point in time, the neural structure was still incredibly simple and could even be integrated into traditional clients with very little overhead.)

But to get it to actually interact, I overcame this by adding a slight reward for increasing distance from 0,0. It is nothing compared to the penalty for failing to survive, but it causes the player to actually move.

There is a key difference here between inputs to the network, and inputs to the training system. The training system decides when the network deserves a reward or punishment (for example, if the player has just respawned, this is detected by the training system and results in a strong punishment against the network’s behavior leading up to the event). The policy network just sees the screen and the world around it. The health of the player is actually ignored by both halves by default. The network sees the health bar, but it doesn’t fundamentally know what it means, since it isn’t punished for decreasing it. But it *learned* over time that decreasing those little red hearts commonly leads to huge punishments (dying). The exact same thing applied to the hunger bar. It had to learn that low hunger started to correlate with dying, then to learn that eating food helped prevent that.

At this point, I realized I had to stop tinkering with it and putting in my own biases. Perhaps it would find a way to survive near spawn. I decided to remove any reward for going away from 0,0 in the nether, and remove any reward for going more than 2000 blocks in the overworld. From then on, it was on its own.

I was shocked with what it came up with after that. Initially, it learned how to attack and kill passive mobs for food, almost immediately upon seeing them. This was probably a holdover from the “learning by example” phase. It still “knows” all of the techniques it learned from there, but now the reinforcement learning neural network is in control, and makes the decisions of what to use when. For example, it’s possible that in the first phase of training it learned that when the hunger bar appears low, a human player will commonly kill a passive mob if it sees one. ****Only now, it’s starting to understand why.**** Essentially, this refinement stage is the network figuring out which techniques it learned in the first phase are actually applicable to its goal.

This brought the training to an interesting point. It knew how to get out of spawn, recognize pig spawners placed in the world, recognize passive mobs. At this point, a standard reinforcement learning algorithm would stagnate. It has achieved everything it gets rewarded for, and avoided everything it gets punished for. It doesn’t get AFK kicked, it doesn’t die of hunger.

But this isn’t a standard reinforcement learning algorithm. This has been purpose designed for Minecraft by researchers out of Carnegie Mellon to properly understand hierarchical goals and sparse rewards. Yet this step was particularly challenging because there’s no reward, just a lack of punishment.

Eventually, it overcame this plateau and started to kill mobs.

This was an absolutely key step, but thanks to the first phase of imitation learning and the high quality input dataset showing combat, the bot was able to pick up on this pattern and recognize how to defend itself once attacked. Previously, it had avoided mobs and even darker cave areas because it associated them with a chance of dying. But now, it sees them as targets to preemptively take out. Skeletons were key in this step of learning. You can't just ignore or avoid skeletons, or wait for them to come to you. You have to actually go after them in order to eliminate them in order to avoid dying.

It eventually learned to transfer this knowledge of skeletons to all mobs, recognizing that all are a threat but ignoring ones far enough away that they aren't relevant.

At this point, I let it run overnight for a few days and checked back.

Amazingly, it had learned to navigate the incredibly sparse reward structure of going into the nether, navigating away from 0,0, and returning to the overworld. If you recall, I had given it a slow reward building up over time for escaping from 0,0 in the overworld, but no reward for going through the nether. As it turned out, the way I implemented this resulted in the bot "building up" all that reward while going through the nether, and getting at all at once upon returning to the overworld much farther away from spawn. It's easy in hindsight to see why this happened: the same reward can be earned 8x faster just by walking through the nether to escape spawn.

Honestly, I believe it randomly learned this behavior by one of the generations of network having a random propensity to walking towards purple, which resulted in emergent behavior of entering then exiting a nether portal.

But, the nether is treacherous. There are players on the highways and at 0,0, and there are gaps in every highway, obviously. Unlike the overworld, there's no chance of a soft landing.

It took some time for the bot to realize what techniques it needed to brave the nether.

But, as a simple agent trying to maximize its reward over time, it was able to take a success rate that was approximately 1/4th its overworld success rate, in exchange for achieving its reward 8 times faster.

It was at this point that it learned more advanced skills almost bordering on "communication". Zombie pigmen are just another mob and potential threat to it, but as I had made sure that each mob was encoded differently into its neural inputs, it was able to eventually learn that they operate by different "rules", in that they only attack when provoked, but when they attack you, you must fight back to have a chance of survival. I don't think the bot would have been able to learn this if it hadn't previously mastered dealing with skeletons, just as it wouldn't have been able to deal with skeletons without having previously dealt with zombies, creepers, and spiders. Shockingly quickly, the bot was able to learn to ignore zombie pigmen, not just not attack and avoid, but entirely ignore, until they attacked it. For a time, it hit the false positive of believing that a zombie pigman oriented directly towards it was a threat, and attacking, but eventually it learned to be rotation agnostic, and recognize when a zombie pigman is actually attacking.

Treating zombie pigmen as a potential threat, but not a certain one, resulted in the success rate nearly doubling (but still less than half of its overworld success rate).

After it learned to survive, I needed to figure out how to harness it. It's a great autonomous bot to play the game, but we don't want it to play the game. We want it to do our bidding.

To accomplish this, I needed to look at how the bot coordinates of all these different learned tasks and goals. This is accomplished through an internal state that is an integral part of the bot's policy neural network. It uses what's called a LSTM (long short term memory) to keep track of the world around it. It doesn't store any of the bot's built up knowledge of how to interact with the world, but rather its current knowledge about the scenario it's in at the moment. For example, things like "I just saw a nether portal over that hill to the right, so I'm going that way to get to it instead of straight away from spawn".

I can influence this internal state to be whatever I want, by using backpropagation to figure out what internal neural structures cause certain actions.

By taking a snapshot of the neural states before and after aggressively attacking the bot but still before it retaliates, it's possible to determine exactly how to influence it to attack another player, for example. You can think of this like analyzing the brain of an alien species in a fMRI: we need to see what situations activate different parts of the "brain" to see what is responsible for what.

I simply artificially induce those neurons I located in the previous step to activate whenever the "attack" command is given, on whatever target is desired.

Previously, one person could only really control one account on 2b. Sure you could open multiple windows to wait through queue, with anti afk on an alt. But there was no way to directly control multiple accounts truly simultaneously (not switching between accounts every few minutes). But with this automation of PvP, one person can be represented by an entire swarm of bots that make into reality whatever they want to happen on the server.

My next avenue of research will be into teaching them to build more complicated structures. While they do build hidey holes sometimes, there is some very interesting research from Facebook AI that just came out in late July that I'm looking at integrating: <https://research.facebook.com/wp-content/uploads/2019/07/CraftAssist-A-Framework-for-Dialogue-enabled-Interactive-Agents-v3.pdf>



Facebook's AI team focused more on how to process English natural language communication with other players to collaboratively play Minecraft. This paper is really more of an advance in language processing than in playing survival minecraft (e.g. their bots use creative mode to build). But I'm confident that when combined with my gameplay networks the result will be basically equivalent in survival.

<https://ai.facebook.com/blog/craftassist-platform-for-collaborative-minecraft-bots/> (new)

<https://www.microsoft.com/en-us/research/academic-program/collaborative-ai-challenge/> (old)

I'm currently neck deep in mashing these together, craftassist on top of my modified minerl, but if/when it's complete, the bots will be able to communicate in English over /msg. The tricky part is retraining craftassist to express its desires in a way that our policy network can understand and carry out.

I'm sending along a zip with a bunch of footage from training. As promised, I've also attached the footage from one of the bots that attacked xdtfwew. As you can see, this is neither a player nor Baritone. It took me a while to put together a visualization from the neural logs, but I'm happy with the results. On the left you can see areas of the video that have greater activation, with different colors corresponding to how they eventually influence the policy network. There are overstimulation issues in close quarters that I'm trying to work out, as you can see. The text on the bottom left is the result of my modal classifier which is predicting the behavioral mode from the internal policy network state.

<https://www.youtube.com/watch?v=S02BHmWPZNs>



FitMC Yesterday at 7:44 AM

I'm looking at some of this footage, it's very impressive
I have a lot of questions but I'm going to compile them into a list
But this is very overwhelming, not in a bad way of course, haha

FitMCToday at 1:08 PM

Have the bots had any experience with crystal pvp yet?

Some; let me explain.

Obviously, crystal pvp is the current meta for the most part. The bots can handle 32ks just as easily as normal swords once I added that in, but those are likely being patched soon. So I've been trying for quite some time to get them to use crystals, but there are various challenges. I haven't really tried in the last few weeks since I'm waiting to see if I can come up with any new ideas to throw at it.

First, the training data and resource issue. The bots effectively need to be handed crystals (whereas in their training data it is quite easy to punch a tree and make a wooden sword to immediately train on how to use it). I overcame this by standardizing a kit shared amongst all. Second, the orientation and looking issue. With a sword, you need to look directly at your opponent to have any effect. This is the cause of the mid-range view locking in the current network (which you could see in the bot video I sent). I had to completely unlock the view direction relative to the target to allow for crystal placement. This causes strange behavior since the network up until that point did not have to think about which direction to face, it was decided for it.

Third, the sparsity of opportunities. Crystals can only be placed on obsidian or bedrock, with empty space above not occupied by any entities. Basically the entire original MineRL dataset needs to be thrown out, since it is almost entirely overworld and doesn't apply at all. There is no such issue with swords, and the MineRL dataset has hundreds of thousands of usable examples of sword combat.

Fourth, the collateral damage issue. I ran into this early on with sword combat, but it was easily solved by removing the "sweeping edge" sword enchant from their kits. The possibilities for collateral damage are just too high with crystal pvp so I've gone back to just one bot running at a time so as to learn the basics, sadly. Given the option to sword or crystal pvp, they would too often opt for swords because, remember, they are a single network and a death of any member is a punishment for all. Given only crystals they would opt to run away. This is why I've been training them solo. I'm considering going back to a swarm, but simultaneously decreasing the punishment for collateral damage just to get them to try it and self improve. At the same time,

this is not as much of an issue as it seems. In a swarm, sword "recharge" can be shared amongst the bots as they take turns. But crystals are not speed limited at all, and you can do damage faster than the invuln clears up after being hit. So only one bot might actually not be too much of a handicap.

I'll dust off the solo crystal network and stage an altercation between me and a bot (on constantiam), give me some time.

Fit:

So while they currently don't have very much experience with crystal pvp, could it be possible for a bot to eventually be on par with a crystal pvp'er, assuming they had access to the standard kit you mentioned?

I'm sure it would take a lot of time, work, and patience

Conner:

It's just a matter of training data and time to make a solo bot crystal pvp, but as I explained it will take a significant amount of development work and some novel approaches to make a swarm that can crystal pvp as one, without tripping each other up. I may have to do another deep dive into the network to manually tweak the neural weights, with the goal being to "surgically" remove this crippling fear of accidentally killing a fellow swarm member

I'm still going to try and get you some footage of the solo crystal network as it exists btw



FitMC Today at 10:00 AM

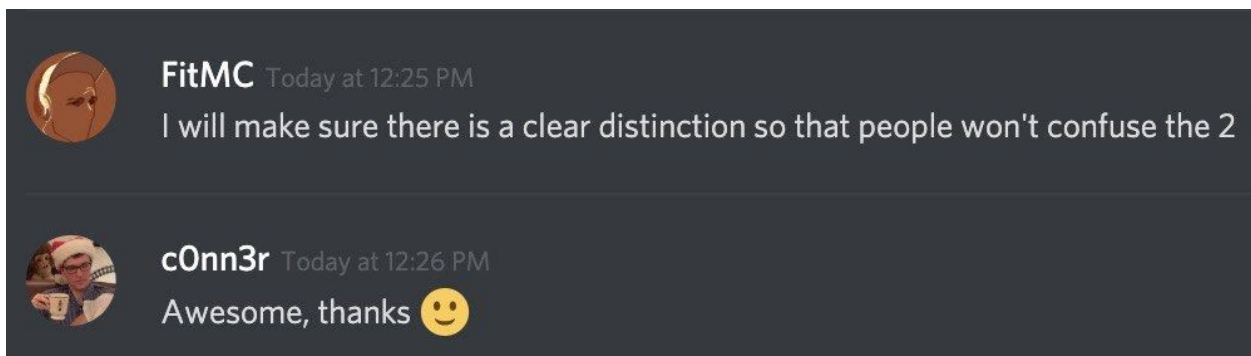
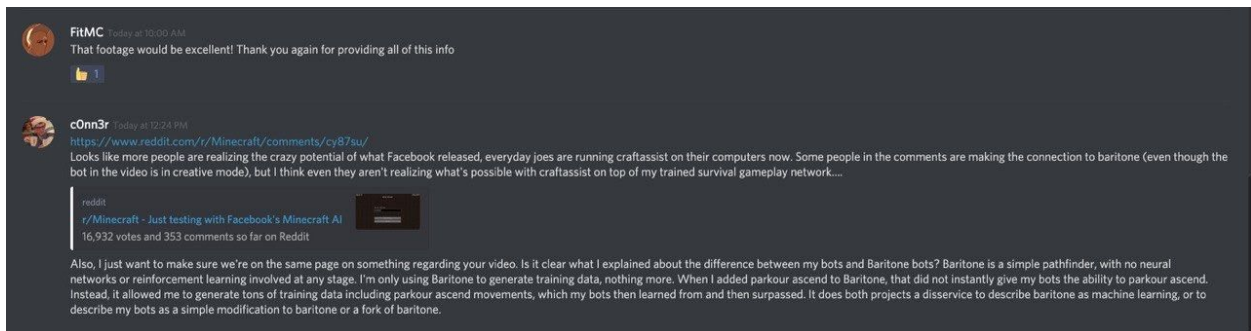
That footage would be excellent! Thank you again for providing all of this info

<https://www.reddit.com/r/Minecraft/comments/cy87su/>

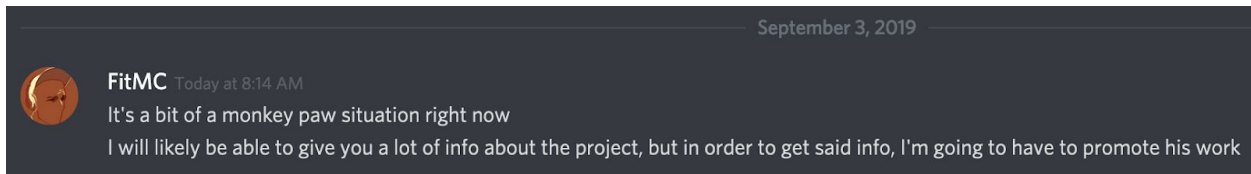
Looks like more people are realizing the crazy potential of what Facebook released, everyday joes are running craftassist on their computers now. Some people in the comments are making the connection to baritone (even though the bot in the video is in creative mode), but I think even they aren't realizing what's possible with craftassist on top of my trained survival gameplay network....

Also, I just want to make sure we're on the same page on something regarding your video. Is it clear what I explained about the difference between my bots and baritone bots? Baritone is a simple pathfinder, with no neural networks or reinforcement learning involved at any stage. I'm only using Baritone to generate training data, nothing more. When I added parkour ascend to baritone, that did not instantly give my bots the ability to parkour ascend. Instead, it allowed me to generate tons of training data including parkour ascend movements, which my bots then learned from and then surpassed. It does both projects a disservice to describe baritone as

machine learning, or to describe my bots as a simple modification to baritone or a fork of baritone.



Separately to leijurv



My response:

we haven't talked at all since, any info would be great at this point. i'll have to trust your judgement on that tradeoff then...

<https://twitter.com/StephanZheng/status/1168945311894687746>

i saw this tweet, seems like the people conner used to work with at salesforce are now releasing papers on minecraft too? i wonder if there's any connection there, has conner mentioned anything like this?

Unsure how to fake the crystal footage. That long message was good cover for only one bot with more fluid motion.

Elon_Musk will be the bot and septave will be the target.

It needs to be on constantiam.

Conner said he would personally be the victim, so we can't have anything like conner being in control of the bot, since he said he would stage it and the narrative is that he's working alone.

Bot needs to look 100% autonomous therefore.

Instead of baritone + kill aura, this will be an actual player controlling the bot.

Perhaps modules to lock your look to the feet of the closest entity would be helpful.

Smooth camera maybe?

Or a less aggressive yaw lock

Talked through it with infernopickaxe and came up with these ideas ^ but he didn't have time to record it with me sadly

Recorded with 0x instead, benefit is the "conner" pov is future on literally retarded settings (left handed LMAO)

<https://www.dropbox.com/s/4kg7c87l2g6tho0/2019-09-04%2000-30-04.mp4?dl=0>

Faked it but it's shit

How do we explain this?!??!

Like so:

(these two links are the same as the youtube links, just source quality)

<https://www.dropbox.com/s/nhcnda526qn1y3j/2019-09-04%2014.14.20.mp4?dl=0>

<https://www.dropbox.com/s/6y9da7wfox4tlz6/my%20pov%20trimmed.mp4?dl=0>

<https://youtu.be/GnSvtlnyt3U>


<https://youtu.be/-do65OneqNs>

First, note that while I did kill the bot, this was not entirely at its fault. I forgot that this old network had never learned totem placement in the offhand. Specifically, note how first the bot popped one of my totems, I popped the bot's totem, it popped my second totem, _then_ I killed it. That would have popped its second totem, if the totem had been replaced. So the bot was actually winning the fight, it was doing damage to me far faster than I was to it. Conditions were also quite unfavorable: fluctuating tps, and the network at this point (just under a month ago) hadn't yet been fine tuned for constantiam's anticheat, so it got rubberbanded a handful of times. (there is no punishment applied for rubberbanding, the network only avoids it as it knows it delays whatever it was trying to do).


The footage cuts off in the moment of death because this puts the network into training mode instead of evaluation mode, so the logs end exactly there. This causes it to backpropagate, accept the punishment, and revise its behavior for the future. It doesn't know exactly what it did wrong, but over the course of literally hundreds of thousands of examples, it starts to understand how to minimize punishment. Also, the drifting in the initial part of the video is that reluctance I mentioned earlier to enter a situation that commonly leads to large punishments.

c0nn3r Today at 2:12 PM
<https://youtu.be/Gn5vtlny13U>
<https://youtu.be/-do65OneqNs>
First, note that while I did kill the bot, this was not entirely at its fault. I forgot that this old network had never learned totem placement in the offhand. Specifically, note how first the bot popped one of my totems, I popped the bot's totem, it popped my second totem, *then* I killed it. That would have popped its second totem, if the totem had been replaced. So the bot was actually winning the fight, it was doing damage to me far faster than I was to it. Conditions were also quite unfavorable: fluctuating tps, and the network at this point (just under a month ago) hadn't yet been fine tuned for constantiam's anticheat, so it got rubberbanded a handful of times. (there is no punishment applied for rubberbanding, the network only avoids it as it knows it delays whatever it was trying to do).
The footage cuts off in the moment of death because this puts the network into training mode instead of evaluation mode, so the logs end exactly there. This causes it to backpropagate, accept the punishment, and revise its behavior for the future. It doesn't know exactly what it did wrong, but over the course of literally hundreds of thousands of examples, it starts to understand how to minimize punishment. Also, the drifting in the initial part of the video is that reluctance I mentioned earlier to enter a situation that commonly leads to large punishments.

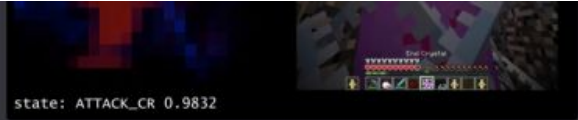
YouTube
Conner Vercellino
ATTACK_CR Preliminary Network Demo, My PoV - constantiam.net
- Sep...



YouTube
Conner Vercellino
ATTACK_CR Preliminary Network Demo, Bot PoV - constantiam.net
- Se...



state: ATTACK_CR 0.9832



state: ATTACK_CR 0.9832

FitMC Today at 2:28 PM
That footage is incredible
I'm still blown away by how far you've advanced this thing
So many people are about to be blown the fuck out by just that clip

c0nn3r Today at 2:29 PM
Thanks 😊

FitMC Today at 2:30 PM
Anyway, progress is coming along. Almost finished with the script. Once that's done, the video usually only takes one extra day
So we're still on schedule for release on Saturday at 10am EST

c0nn3r Today at 2:30 PM
Great, super excited!

HOLY FUCKING SHIT

So, now that I've done my part of providing you all this information, I'm curious how you're holding up your end. I know the material I provided is pretty extensive, let me know when you

come up with that list of questions. Not to be blunt, but I want to make sure you don't look like an idiot trying to explain this stuff...

What, if any, of this content will you cut and not include in your showcase video? How long will it be approximately? Any chance you'd be willing to share the script? We both want this to be the best it can be!

^ Sent to fit after some deliberation. Upside is basically the alignment of incentives. Fit really doesn't want to look like an idiot (see all his "i told you so"s). Even if fit was embarrassed to ask some dumb questions this should make him unstuck there lol. He also literally said he would make a list of questions so that's what that refers to. Slight downside is rocking the boat I guess. Fit might take it the wrong way? Herm says that this will work, so we did it XD.

FIT SENT A SCRIPT BUT ITS SHIT SO WE EDITED IT

The Man Behind 2b2t's Robot Invasion

Are you going to include the footage of the bot PoV of xdtfwew's altercation? It could provide immediate credibility that the person who you're talking to is truly the one behind it all, hooking in viewers who were interested in your previous video.

(crystal pvp clip opens)

Overall I think your decisions on what parts to include and what parts to cut are a little strange, and make it a little nonsensical. I've tried to add back some sections that I think are key.

I've essentially only made additions (bold), but the two sentences I deleted I put in strikethrough.

Here is my explanation for why I added what I did:

- **First and foremost, as a showcase for my work, I have no higher priority than the section that explains my innovation in creating these bots: three phases of training, getting more and more specialized, and augmenting a human recorded dataset with a simpler bot recorded dataset (using Baritone in this case). I added a paragraph or two on this. It isn't really a good demonstration if it doesn't actually explain what I did differently from MineRL to make me succeed where they failed.**

- **Second priority is explaining exactly what rewards and punishments I gave it, but you got almost all of that correct (reward for getting away from 0,0, punishment for death, punishment for kick). So I just added a small sentence.**
- **I would like if you would mention the MineRL project, since they were truly instrumental. As the saying goes, I was standing on the shoulders of giants. I added two references to them. You can link their site in the description or not, I don't particularly mind since anyone who is interested can google them.**
- **Clarifying the difference this project and Baritone is also important, but you've already done this nicely.**
- **Explaining my future plans regarding craftassist is not necessary, but since you explicitly asked for my future plans, I've added it at the bottom. I've also expanded more on the monetization issue, since this has become more and more pressing in the last week.**
- **I'm not sure why you removed the history with leijurv and his hypocrisy, but if you don't think it's interesting then fine.**

What you are seeing here is nothing short of incredible. An a.i. attempting the most sophisticated form of combat in Minecraft, crystal pvp.

6 months ago, this feat was said to be impossible, but through hard work, patience, and dedication, one man has been able to accomplish the impossible.

He goes by the name of "c0nn3r", and he's been conducting his a.i. tests and experiments for the past few months here on the **oldest anarchy server in minecraft? :)**

I've been in touch with him for about the past month or so, and he's shared some insight into:

-How his bots work

-**Why** he chose 2b2t (and other anarchy servers) for conducting his experiments.

-His future plans for the server

We have a lot to discuss today, so let's get started.

Before we begin, I have to point out that in my previous videos, I stated that c0nn3r's bots were part of the baritone project. While baritone's capabilities are excellent, conner's work is completely separate from baritone, and thus, is not currently available to the public.

He gave me an **exclusive** behind-the-scenes look at how his bots were able to learn, train themselves, and perfect their Minecraft gameplay. The history behind these bots is quite fascinating. He went into very technical details but I've done my best to simplify these details for you to understand.

Conner's bots learned how to play the game in **three** phases. Each phase being more technically demanding than the last, **requiring larger and larger neural networks and more and more training time.**

In their first two phases, the bots learned from recorded Minecraft gameplay. This is based on an existing project called MineRL who collected this data and released it in April. Conner's innovation was to use Baritone to generate thousands of hours of skilled gameplay footage very quickly and efficiently, running the game at a speed far faster than a human's reaction time. This is what led to the confusion about the relationship between Conner's bots and Baritone.

In their first phase, the bots trained themselves to replicate the basic actions **to play Minecraft. This is from that combined dataset of sixty million human examples (from MineRL) and hundreds of millions of examples from Baritone. Each "example" takes the rough form of "in *this* situation, while working towards *this* goal, a human player took *this* action, such as 'pressed W while moving the camera to the left slightly'".**

In the second phase, they specialized for anarchy, being fed in footage from 2b2t gameplay, bedrock level pathing, et cetera. Conner imported bulk footage from 2b2t streamers such as myself exploring the greater spawn area. This allowed them to learn the actions you would perform while exploring spawn, while building, etc. ~~The bots were shown hours and hours of minecraft gameplay and also directly fed information.~~

~~They were able to quickly and easily learn simple facts about Minecraft and the server specifically.~~

Then, in their **third** phase, they actually took control of a player in game, now that it's learned what to do. This stage is called "refinement". It plays through Minecraft over and over, but, **instead of trying to replicate the decisions made by a human or Baritone, it is now trying to make whatever decisions are necessary to achieve its goal, which on 2b2t, is to survive. Every time the bot dies, it considers death a "punishment", and with enough deaths, it will begin to actively avoid this "punishment" however it can, by trying to survive for as long as possible. A secondary much smaller reward was added to nudge it away from just standing stationary at spawn, and a comparably sized punishment was added for being AFK kicked.**

At first, the bot ignored the health bar and hunger bar. **It had no inbuilt knowledge that that area of the screen was even important.** But it learned over time that decreasing those little red hearts commonly leads to huge punishments (dying). It had to learn that low hunger correlated with dying, and that eating food helped prevent that.

Conner decided to tweak the bots so that they would not **receive any further rewards for going** farther away than **a few thousand** blocks from 0,0 so they could continue to learn in a very hostile environment.

He was shocked with what it came up with after that. Initially, it learned how to attack and kill passive mobs for food, almost immediately upon seeing them. It likely remembered seeing other players kill animals for food if their health became low during **stages one and two of training.**

In general, you can think of this refinement stage as the network figuring out which techniques it learned in the first phase are actually applicable to its goal, and which can be discarded.

This brought the training to an interesting point. It knew how to get out of spawn, recognize pig spawners placed in the world, and recognize passive mobs. At this point, a standard reinforcement learning algorithm would stagnate. It has achieved everything it gets rewarded for, and avoided everything it gets punished for. It doesn't get AFK kicked, it doesn't die of hunger.

Given the buildup of "At this point, a standard reinforcement learning algorithm would stagnate" I think you need the payoff of something like this paragraph, but I understand if you have to reword it:

But this isn't a standard reinforcement learning algorithm. This has been purpose designed for Minecraft by researchers out of Carnegie Mellon to properly understand hierarchical goals and sparse rewards. Yet this step was particularly challenging because there's no reward, just a lack of punishment.

Eventually, it overcame this plateau and started to kill mobs.

This was an absolutely key step. Previously, it had avoided mobs and even darker cave areas because it associated them with a chance of dying. But now, it saw them as targets to preemptively take out. Skeletons were key in this step of learning. You can't just ignore or avoid skeletons, or wait for them to come to you. You have to actually go after them in order to eliminate them.

It eventually learned to transfer this knowledge of skeletons to all mobs, recognizing that all are a threat but ignoring ones far enough away that they aren't relevant.

At this point, conner let the program run for a few days and checked back.

Amazingly, it had learned to navigate by going into the nether, getting away from 0,0, and then returning to the overworld.

It **was** easy in hindsight to see why this happened: the same reward can be earned 8x faster just by walking through the nether to escape spawn.

Conner believes it randomly learned this behavior by one of the generations of network having a random propensity to walking towards the color purple, which resulted in emergent behavior of entering then exiting a nether portal.

But, the nether is treacherous. There are players on the highways and at 0,0, and there are gaps in every highway, obviously. Unlike the overworld, there's no chance of a soft landing.

It took some time for the bot to realize what techniques it needed to brave the nether.

But, as a simple agent trying to maximize its reward over time, it was able to take a success rate that was approximately 1/4th its overworld success rate, in exchange for achieving its reward 8 times faster.

It was at this point that it learned more advanced skills almost bordering on “communication”. Zombie pigmen are just another mob and potential threat to it, but as conner had made sure that each mob was encoded differently into its neural inputs, it was able to eventually learn that they operate by different “rules”, in that they only attack when provoked, but when they attack you, you must fight back to have a chance of survival. The bot wouldn't have been able to learn this if it hadn't previously mastered dealing with skeletons, just as it wouldn't have been able to deal with skeletons without having previously dealt with zombies, creepers, and spiders. Shockingly quickly, the bot was able to learn to ignore zombie pigmen, not just not attack and avoid, but entirely ignore, until they attacked it. For a time, it hit the false positive of believing that a zombie pigman oriented directly towards it was a threat, and attacking, but eventually it learned to be rotation agnostic, and recognize when a zombie pigman is actually attacking. Treating zombie pigmen as a potential threat, but not a certain one, resulted in the success rate nearly doubling.

It seems like you've cut the part where I explain how I harness these bots under my control, I'm curious why? Is it too complicated? It's a key step that brings it from an autonomous agent to one that's controlled by its owner.

By this point, the bots had become perfectly capable of thriving on 2b2t. From here, conner explained his plans for teaching the bots how to properly build structures. They had learned to build small hidey holes to protect themselves, but no actual structures yet.

Include the section on crystal pvp and why it's hard? You don't mention crystal pvp at all in the script yet, and I think the viewers would be very curious how this part works, since it was the hook at the beginning. It should give them an appreciation for how impressive the footage you're showing truly is.

Obviously, crystal pvp is the current meta for the most part. The bots can handle 32ks just as easily as normal swords once I added that in, but those are likely being patched soon. So I've been trying for quite some time to get them to use crystals, but there are various challenges. I haven't really tried in the last few weeks since I'm waiting to see if I can come up with any new ideas to throw at it.

First, the training data and resource issue. The bots effectively need to be handed crystals (whereas in their training data it is quite easy to punch a tree and make a wooden sword to immediately train on how to use it). I overcame this by standardizing a kit shared amongst all.

Second, the orientation and looking issue. With a sword, you need to look directly at your opponent to have any effect. This is the cause of the mid-range view locking in the current network (which you could see in the bot video I sent). I had to completely unlock the view direction relative to the target to allow for crystal placement. This causes strange behavior since the network up until that point did not have to think about which direction to face, it was decided for it.

Third, the sparsity of opportunities. Crystals can only be placed on obsidian or bedrock, with empty space above not occupied by any entities. Basically the entire original

MineRL dataset needs to be thrown out, since it is almost entirely overworld and doesn't apply at all. There is no such issue with swords, and the MineRL dataset has hundreds of thousands of usable examples of sword combat.

Fourth, the collateral damage issue. I ran into this early on with sword combat, but it was easily solved by removing the "sweeping edge" sword enchant from their kits. The possibilities for collateral damage are just too high with crystal pvp so I've gone back to just one bot running at a time so as to learn the basics, sadly. Given the option to sword or crystal pvp, they would too often opt for swords because, remember, they are a single network and a death of any member is a punishment for all. Given only crystals they would opt to run away. This is why I've been training them solo. I'm considering going back to a swarm, but simultaneously decreasing the punishment for collateral damage just to get them to try it and self improve. At the same time, this is not as much of an issue as it seems. In a swarm, sword "recharge" can be shared amongst the bots as they take turns. But crystals are not speed limited at all, and you can do damage faster than the invuln clears up after being hit. So only one bot might actually not be too much of a handicap.

It will take a significant amount of development work and some novel approaches to make a swarm that can crystal pvp as one, without tripping each other up. I may have to do another deep dive into the network to manually tweak the neural weights, with the goal being to "surgically" remove this crippling fear of accidentally killing a fellow swarm member

Include the section about my current work? That answers your third question (my plans for the future). See craftassist and the footage I provided, pasted here:

<https://www.reddit.com/r/Minecraft/comments/cy87su/>

Previously, one person could only really control one account on 2b. Sure you could open multiple windows to wait through queue, with anti afk on an alt. But there was no way to directly control multiple accounts truly simultaneously (not switching between accounts every few minutes). But with this automation of PvP, one person can be represented by an entire swarm of bots that make into reality whatever they want to happen on the server.

My next avenue of research will be into teaching them to build more complicated structures. While they do build hidey holes sometimes, there is some very interesting research from Facebook AI that just came out in late July that I'm looking at integrating:
<https://research.facebook.com/wp-content/uploads/2019/07/CraftAssist-A-Framework-for-Dialogue-enabled-Interactive-Agents-v3.pdf>



Facebook's AI team focused more on how to process English natural language communication with other players to collaboratively play Minecraft. This paper is really more of an advance in language processing than in playing survival minecraft (e.g. their bots use creative mode to build). But I'm confident that when combined with my gameplay networks the result will be basically equivalent in survival.

*<https://ai.facebook.com/blog/craftassist-platform-for-collaborative-minecraft-bots/> (new)
<https://www.microsoft.com/en-us/research/academic-program/collaborative-ai-challenge/>
(old)*

I'm currently neck deep in mashing these together, craftassist on top of my modified minerl, but if/when it's complete, the bots will be able to communicate in English over /msg. The tricky part is retraining craftassist to express its desires in a way that our policy network can understand and carry out.

For "why 2b2t" I always say 2b2t is the only environment that can provide something hard enough for these bots, densely populated with other intelligent yet adversarial actors with which it could communicate. What I mean by this is that other Minecraft servers are too tame. People will help the bots too often, they rarely get killed by intelligent actors. Instead of learning true evasion of threats, it just learns to evade simple Minecraft mob pathfinding. The initial learning in phase 1 is just from vanilla Minecraft, I experimented with allowing it to refine itself on vanilla environments but it just so rarely got an actual punishment that it didn't really work. It just isn't hard enough.

This is why I added a phase 2 which was anarchy specialized footage, as a stepping stone to stage 3 which is autonomous gameplay in anarchy. Overall, 2b2t and other anarchy servers offer a profoundly challenging environment, and results in a bot that can truly handle anything that Minecraft has to throw at it.

I plan to sell this bot to the highest bidders. The hardware required to train even 5 of these bots is expensive and I require a constant source of funding. Each one requires an Nvidia P100, a nearly \$6k GPU. I am renting them from Microsoft Azure and it is beginning to become unsustainable, I only train in short bursts now. Running the model after it's been trained can be done on a consumer grade GPU (but certainly not on integrated graphics). Refer to the screenshot I sent you last week of my explanation of how I want to monetize: (even though this was written some time ago for baritone, it applies just as well for my bot swarm)

<https://cdn.discordapp.com/attachments/611205819224293395/614578998944464897/monetization.png>

September 5, 2019



FitMC Today at 3:02 AM

When the script is finished in a few hours I'll post it here so you can take a look at it




FitMC Today at 3:22 AM

So here's the 3 main questions that I had:
-How do your bots work? (You've already answered this)
-Why did you choose 2b2t for your experiments?
-What are your future plans on the server?



FitMC Today at 8:17 AM

This is about the first 7 minutes of the video. It's still a rough copy so if you see something that you don't like, I can edit or remove it.

 [conner1.txt](#)
6.35 KB



It brings the viewers up to the point where you speak about wanting to teach the bots structures
After that, I want to touch on how the work on crystal combat is still ongoing, but then get into questions #2 and #3
I'll be back at my desk in about 2 hours and then I'll be there for the rest of the day



c0nn3r Today at 10:43 AM

Working on some improvements and answering your questions right now



FitMC Today at 11:00 AM

Sounds good



c0nn3r Today at 11:39 AM

Should be done with my suggestions within half an hour.

So far it's at 9:35 minutes, so now I will type out the section about crystal pvp
Then finish with the responses you gave to the last two questions



c0nn3r Today at 1:17 PM

Epic.

At the end, if you go into the monetization disagreement between me and lejjury (part of my answer to the third question), perhaps mention what I explained in the initial message (lejjury's hypocrisy regarding the "bots" video <https://www.youtube.com/watch?v=uQXhba1lVOI> and him killing off the feature last year: <https://discordapp.com/channels/@me/611205819224293395/614579027939688491>) (edited)

YouTube

Lurf Jurv

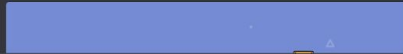
bots



Discord

[Discord - Free voice and text chat for gamers](#)

Step up your game with a modern voice & text chat app. Crystal clear voice, multiple server and channel support, mobile apps, and more. Get your free server now!



FiMC Today at 1:56 PM

I think you said it best when you told me about crystal pvp so I'll be heavily quoting you
Here's what I have:

Now, here is where things get absolutely insane. When it came time to teach the bots crystal pvp, Conner gave me some insight into just how much more difficult it was to teach them something so sophisticated.

"I've been trying for quite some time to get them to use crystals, but there are various challenges. First, the training data and resource issue. The bots effectively need to be handed crystals (whereas in their training data it is quite easy to punch a tree and make a wooden sword to immediately train on how to use it). I overcame this by standardizing a kit shared amongst all.

Second, the orientation and looking issue. With a sword, you need to look directly at your opponent to have any effect. This is the cause of the mid-range view locking in the current network (which you could see in the bot video I sent). I had to completely unlock the view direction relative to the target to allow for crystal placement. This causes strange behavior since the network up until that point did not have to think about which direction to face, it was decided for it.

Third, the sparsity of opportunities. Crystals can only be placed on obsidian or bedrock, with empty space above not occupied by any entities. Basically the entire original MineRL dataset needs to be thrown out, since it is almost entirely overworld and doesn't apply at all. There is no such issue with swords, and the MineRL dataset has hundreds of thousands of usable examples of sword combat.

Fourth, the collateral damage issue. I ran into this early on with sword combat, but it was easily solved by removing the "sweeping edge" sword enchant from their kits. The possibilities for collateral damage are just too high with crystal pvp so I've gone back to just one bot running at a time so as to learn the basics, sadly. Given the option to sword or crystal pvp, they would too often opt for swords because, remember, they are a single network and a death of any member is a punishment for all. Given only crystals they would opt to run away. This is why I've been training them solo.

I'm considering going back to a swarm, but simultaneously decreasing the punishment for collateral damage just to get them to try it and self improve. At the same time, this is not as much of an

I'm considering going back to a swarm, but simultaneously decreasing the punishment for collateral damage just to get them to try it and self improve. At the same time, this is not as much of an issue as it seems. In a swarm, sword "recharge" can be shared amongst the bots as they take turns. But crystals are not speed limited at all, and you can do damage faster than the invuln clears up after being hit. So only one bot might actually not be too much of a handicap.

It will take a significant amount of development work and some novel approaches to make a swarm that can crystal pvp as one, without tripping each other up. I may have to do another deep dive into the network to manually tweak the neural weights, with the goal being to "surgically" remove this crippling fear of accidentally killing a fellow swarm member"

—
This video is going to be longer than expected lol, but I'm alright with that 😊



c0nn3r Today at 1:58 PM
Sounds good to me!



FitMC Today at 2:14 PM
I'm done working on it for the day but I won't be adding anything that hasn't already been sent your way. I'll be spending most of the day tomorrow finishing everything and having it ready to go for Saturday



c0nn3r Today at 2:17 PM
I understand if you're hesitant to plug a paid service, but it's my route forwards, so I'd really appreciate if you'd at least touch on or give a brief mention of my monetization plans for the future, just so people are aware that this will eventually be available as a service.
Can't wait for the video, thanks for all of your work on it.



FitMC Today at 2:18 PM

Can't wait for the video, thanks for all of your work on it.



FitMC Today at 2:18 PM

Oh definitely, I'll be quoting you verbatim on the last section

"I plan to sell this bot to the highest bidders. The hardware required to train even 5 of these bots is expensive and I require a constant source of funding. Each one requires an Nvidia P100, a nearly \$6k GPU. I am renting them from Microsoft Azure and it is being to become unsustainable, I only train in short bursts now. Running the model after it's been trained can be done on a consumer grade GPU "

I bet you're going to have a lot of people approaching you after this video



c0nn3r Today at 2:20 PM

Perfect. I'd suggest this as the visual then: <https://cdn.discordapp.com/attachments/611205819224293395/614578998944464897/monetization.png>



Has many of my ideas for this project's future

monetization.png

Has many of my ideas for this project's future



FitMC Today at 2:21 PM



c0nn3r Today at 2:22 PM

Oh last thing, my mistake, it should be "beginning" not "being" (in I am renting them from Microsoft Azure and it is being to become unsustainable) (edited)

Old stuff and resources:

Need to look at <http://miner1.io/assets/videos/> some more and see what we can get
Should we bring up <https://www.twitch.tv/thimen24>

This is a draft of the intro to Fit that goes through all the points, but way too fast. This will not be used except as a reference or to pull ideas out of.

Hey Fit,

Congratulations again on the Baritone video. I'm really excited and inspired by how your fans responded to Baritone. You did a good job explaining such a technical project to your audience and I'm grateful for your assistance in increasing the project's visibility

However, the reason I'm contacting you is to ask you for help. Unfortunately, the very person I asked you to contact (Vercellino) has turned a growing number of my collaborators against me. Baritone is built on MineBot, an earlier attempt at building a pathfinding artificial intelligence. Perhaps because I created Baritone without the permission of other MineBot devs

Vercellino also continued to insist that we train bots to do PVP. I was very against this. I created Baritone because I wanted to make something that could bring people together - not introduce a new type of warfare. Vercellino could not be reasoned with however and began training agents on various anarchy servers. This is why I initially suggested you reach out - I knew he was working on something with huge potential impact to the server.

The minebot developers feel wronged as I created Baritone without their permission. While I had talked to them since we had our falling out they seem to have been looking for an excuse to backstab me. Vercellino pointed out that the technology in Baritone could be used to generate far more revenue than any minecraft project ever [see issue]. I disagreed. While my living conditions are currently non-ideal, I still believe that software should be free and open source. Perhaps because Vercellino comes from a background of sales optimization and corporate R&D, it seems that he doesn't share my ideals. He promised these developers that if they helped him he would share any profits from his Swarm software. He has created a private Baritone copy that has capabilities never seen before in a Minecraft utility mod. I attempted to delete it, but he caught me and is attempting to leverage a copyright takedown to remove Baritone from Github.

I now ask you to bring this to light. I need someone to explain what is coming to the server. To warn everyone. I need someone to help me bring down Vercellino and his ring of conspirators.

Best wishes,

liejurv

My original idea for the first message was hey fit, your baritone vid seems to have been a total success! crazy how many views it got. i was wondering if you might be interested in a follow-up? vercellino has been working on some crazy full-on machine learning extensions to baritone, he won't tell me much but it has to do with natural language processing. he's currently testing primarily on constantiam (due to queue), but it would work on any anarchy server. he says that he will be ready to go public in a week or two, and was wondering if maybe you'd want to make a follow-up video on it. i can put you guys in contact if so.

For context, here is the end of our DM before the video went up (there were a few messages after this where I asked what time it would go up but those dont matter)

leijurv

07/02/2019

actually i know an expert in reinforcement learning and artificial intelligence who works at salesforce and who has advised me on this project, if you want I could connect you up if you want to ask him some questions

leijurv

07/02/2019

if you're interested you can email him at vercellino@mit.edu (i checked and he's more than willing to talk with you about it) or call him at [redacted for this document lol but i did send his actual cell phone]

FitMC

07/02/2019

Thank you for all these resources. I will make sure to link the github page in the description of the video as well as mention it.

leijurv

07/02/2019

Sure thing, glad to help! Let me know if you plan to consult Vercellino for this video, so I can tell him whether to expect you to reach out. I understand if it's too late to add new content, and if so I'm sure he'd be willing to talk for a future video.

FitMC

07/02/2019

I won't be contacting him for anything regarding for this video but **I will keep him in mind for future videos** :thumbup::skin-tone-1: **Curious what he has to say**
(emphasis mine)

The approach:

Leverage craftassist and minerl. There's a huge amount of credibility added by being able to link to things like <https://ai.facebook.com/blog/craftassist-platform-for-collaborative-minecraft-bots/>

Things to remember

- **Fit is retarded and Fit is incredibly suspicious**
 - I sent him the link to the github contributors page and also sent a screenshot of part of it. He presumably didn't want to click a random link and he just showed my screenshot in his video.
 - It is a good idea to send screenshots alongside links, to send full PDF files instead of links to PDF files, I think.
 - He was fine with clicking youtube links tho (he's not 100% retarded just 99%)
 - IronException (who Fit talked to the most for the july 4 vid, before me) said that he tried to get fit to install or run baritone but there was no way Fit would do it
 - He cannot understand what constitutes machine learning or AI
 - He has borne the brunt of the 2b2t community hating him for years and probably trying to hack and rat him the whole time
 - This is both a blessing and a curse: he will not run anything himself, so he will have to trust our footage
- I told Fit that Vercellino had advised me on baritone so far, **we need to come up with something that he helped me with**
 - Perhaps tuning the exact cost values...? (the A* heuristic)
 - Or maybe using reinforcement learning to tune the action costs?
 - It needs to be something that is explainable even though conner has no contributor history. I think that tuning specific parameters to their optimal values fits the bill nicely. For example, we could say that I sent him the source code for baritone and told him which parameters to tune.
 - Playing off of this, what if we have conner reveal like "I told leijurv that I tuned those baritone parameters to get from point A to point B as fast as possible, but really I tuned them to eliminate targets as fast as possible. Sure, part of that included incredibly fast and precise movements, but the goal has always been the ability to hunt down and kill any moving target that's trying to evade us"

- I have publicly stated that there is no machine learning in Baritone. If Fit brings this up, there's an easy cover which is that previous bullet point ^, that conner trained these parameters without my knowledge. I thought he just directly calculated the optimal values, but it was actually a reinforcement learning algorithm.
- Fit will probably ask other players to confirm this, IF we go about it saying that it's public fork of baritone or something like that.
 - There are other people who are actually knowledgeable about baritone's internals who he has a relationship with.
 - This is why the initial message that I send should probably give an **air of secrecy**
 - It's not just that I'm keeping this secret from the community, it's that Vercellino is keeping it secret from me until he's ready
- Fit probably has not seen freecam combined with click. I'm not entirely sure which features he has seen
- The "only goal is to survive" is a play off what fitmc commonly says about the server, but in this case it can actually work.
- **Remember, at the end of his last video he said "as long as these bots don't learn to PvP or think, we'll be okay". We need to make it seem like both of these events have happened**
- Fit has already seen and taken footage from my "bots" video where I had 4 bots controlled by one account in like November or something. We'd have to mention this head-on because it seems at first glance like a contradiction to this narrative. But instead it could be like a "I started down this path, but I realized how dangerous it would be if I perfected and released this. I couldn't see any benefit to me or those I cared about by creating a multi bot control system. Everything I wanted to do, I could do by just controlling one account. Vercellino disagreed with my restraint"

Funny idea: What if we played off of the recency of the craftassist paper and basically said that while the bot has been going for a long while based off of baritone and minerl, we only recently added natural language processing for english. Could say something like: "prior to integration of facebook's AI research regarding a minecraft based assistant that understands english. We used to have the bots communicate to each other by simply sharing neuron activations. Normally, the bots are running from the same computer so we just directly transmitted the neuron values between the different bot windows. But as we scaled up to many computers, we actually ended up using the "/msg" system where the bots would message each other their internal neural state to share information/knowledge about the world. (could show a faked screenshot of a bot literally /msg-ing another bot [-0.4, 0.3, 1.7, -1, 2.4] or some shit like that lol). But now that we've applied the craftassist natural language processing model from facebook, the bots are actually able to express these "thoughts" in english. (another faked screenshot of a bot /msg-ing another bot "watch out, that portal is trapped"). We weren't able to previously do this, because Minecraft is such a narrow domain that existing natural language processing frameworks for machine learning did not properly express concepts that only exist in Minecraft, such as nether portals and end crystals.

Even though minerl says that they're using limited compute resources, in the world of machine learning that's an entirely different scale. They say "For this competition we will specifically be restricting competitors to NC6 v2 Azure instances with 6CPU cores, 112 GiB RAM, 736 GiB SDD, and a single NVIDIA P100 GPU."

<https://research.facebook.com/wp-content/uploads/2019/07/Why-Build-an-Assistant-in-Minecraft-v3.pdf>

nature and their commercial importance makes experimentation with them difficult.

In this work we argue for building an open interactive assistant, and through it, the tools and platform for researching grounded NLU. Instead of a “real world” assistant, we propose working in the sandbox construction game of Minecraft ¹. The constraints of the Minecraft world (e.g. coarse 3-*d* voxel grid, simple physics) and the regularities in the head of the distribution of in-game tasks allow numerous hand-holds for NLU research. Furthermore, since we work in a game environment, [players may enjoy interacting with the assistants as they are developed, yielding a rich resource for human-in-the-loop research.](#)

The rest of this document describes in more depth the motivations and framing of this program, and should

<https://research.fb.com/wp-content/uploads/2019/07/CraftAssist-A-Framework-for-Dialogue-enabled-Interactive-Agents-v3.pdf>

<https://arxiv.org/abs/1907.08584>

<https://github.com/facebookresearch/craftassist>

This readme is crucial

<https://www.microsoft.com/en-us/research/academic-program/collaborative-ai-challenge/>

<https://arxiv.org/pdf/1904.10079.pdf>

I’m not sure if/how minerl competition could be pulled in, but maybe

<http://minerl.io/competition/>

<http://minerl.io/>

These gifs are crucial. We need to make some that show 2b2t footage, ideally even baritone too.

Maybe we could tell him that he can login to minerl and train the agent to navigate terrain.

Idea for how to bring up the idea initially:

Recent events have made me realize that I might need to halt Baritone development.