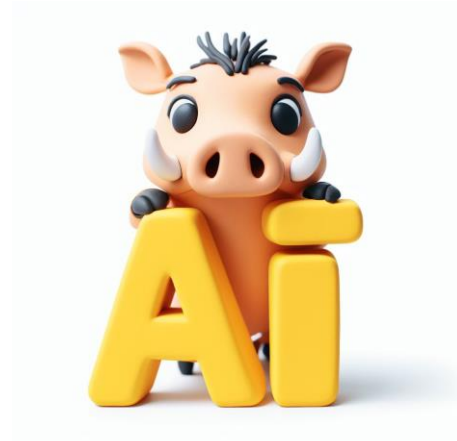# Artificial Intelligence

Lecture 11-3
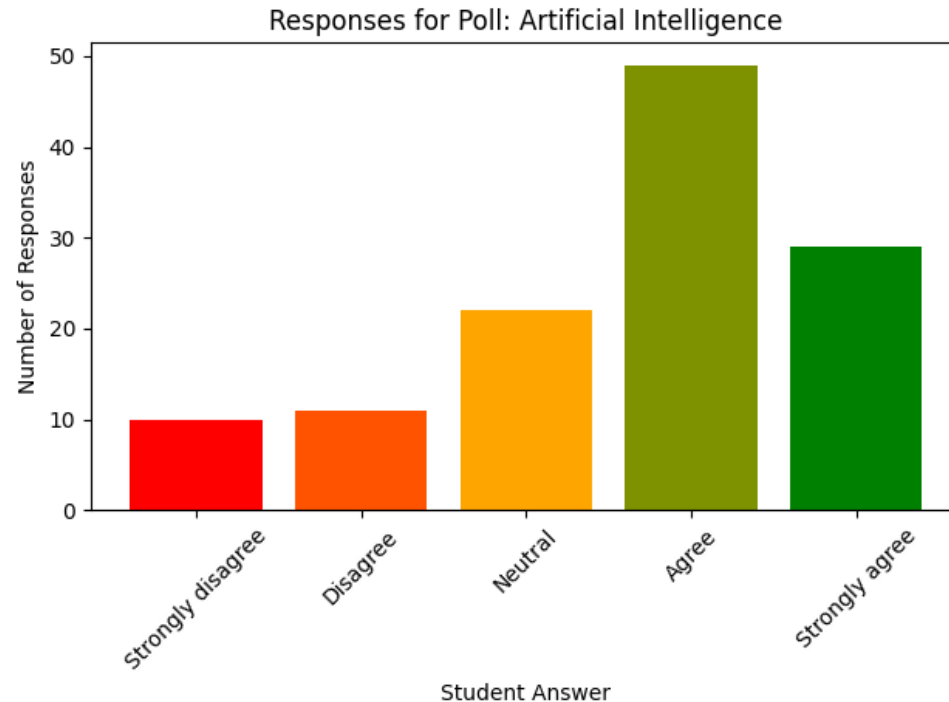
Computers & Society (CPSC 430)

Kevin Leyton-Brown

Concern: Creators

- Will generative AI **harm content creators**?
  - artists, novelists, actors, …

- Is training a generative model on creative work a **violation of copyright law** or "fair use"?
  - should the law change?
  - risk of memorization
  - is it wrong to reference existing artists in style instructions?
    - picture of the Vancouver skyline in the style of Van Gogh's "Starry Night"

- **Photoshop** didn't put photographers out of a job
  - and ultimately helped increase our consumption of photography
  - but is this time different?

Artificial Intelligence

"Content generated by artificial intelligence systems (e.g., ChatGPT, DALL-E) that require minimal human input should not receive copyright protection."
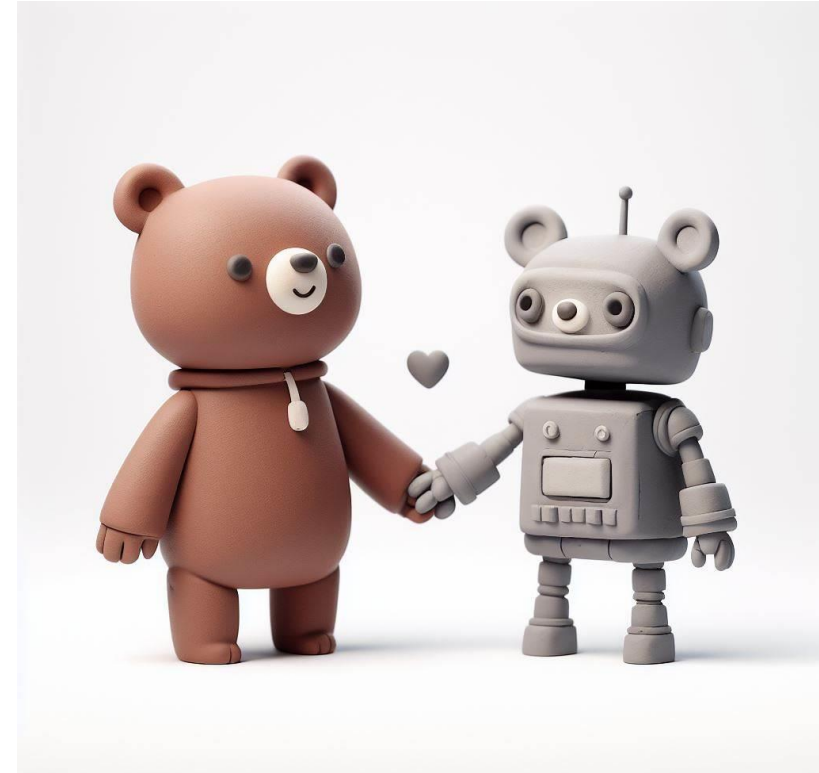

Responses for Poll: Artificial Intelligence

Concern: Education

- ## LLMs will **profoundly affect education**
  - ChatGPT generates decent (~B) essays easily
  - Can also rewrite text, critique arguments,
    solve problem sets, answer comprehension questions

- ## We will need to **reconsider what we teach**
  - is the ability to generate fluent prose still important?
    - analogous to long division when students have calculators
    - many of the same arguments on both sides apply
  - what critical thinking skills become important when
    we assume that students have access to AI?
  - ultimately, our need for education won't go away

- ## We will need to **reimagine evaluation**
  - currently, assignments based on what is hard to do, easy to grade
    - that balance just changed!
  - ultimately, I foresee LLMs serving as part of the solution

Concern: Broader Social Impacts

- How will AI technologies **transform society**?

- Will there be a **social backlash** against AI?
  - If so, what will be considered AI?

- We're raising **children** taking for granted many technologies that strike us as magical

- How will **human relationships** change in the presence of always-available social agents?

- As we are increasingly **augmented by AI**, what are our inherent cognitive/emotional/motivational limitations, beyond which augmentation won't help?

Ethics of AI Will Be a Big Conversation

Will a new technology:

- disempower **individuals vs corporations**?
  - ⇒ user modeling; data mining; fostering addictive behaviors; developmental effects on children

- disempower **individuals vs governments**?
  - ⇒ facilitate disinformation (deep fakes; bots masquerading as people; filter bubbles); enable qualitatively new military or security tactics

- take **autonomous actions** in a way that obscures responsibility
  - ⇒ autonomous weapons; self-driving cars; loan approval systems

- disproportionately affect **vulnerable/marginalized groups**
  - ⇒ automated decision making tools trained in ways that may encode existing biases

Prediction: We'll Get What We Want (for good or bad)

- ## Entertainment
  - gaming is already bigger than Hollywood
  - the lines surrounding the "game" category will continue to blur

- ## Free time
  - automation of routine and unpleasant tasks

- ## A burst of creative energy
  - generative AI will lead to new art forms

- ## Connecting with others
  - technologies for brokering, mediating, facilitating connections between people
  - and/or digital friends and romantic partners (!)

Concern: Superhuman Intelligence

- AI systems are increasingly capable of **human-level performance**

- **Superhuman intelligence** isn't such a foreign, scary thing
  - governments, corporations, NGOs exhibit behavior much more sophisticated and complex than that of any individual

- Many important problems need superhuman intelligence; AI can help
  - improved **collective decision making**
  - more efficient use of **scarce resources**
  - addressing **underserved communities**
  - **climate change**; other societal challenges

Concern: Sentience

- "AI will become **self-aware** and kill us all"

- Today, there's no "**ghost in the machine**"
  - modern AI systems don't form goals or take actions
  - really a mapping from (e.g.) text to more text
  - not sitting there thinking when not in use
  - engineered systems with safeguards
    - ChatGPT can't even tell a dirty joke!

- But **fears persist** about what the future holds
  - rapid rate of improvement, recently accelerating
  - nobody understands the "hard problem of consciousness"
  - effective altruists have made influential arguments about "existential risk"
    - but many of them also say we should focus more on colonizing Mars than e.g. helping poor people alive today

Conclusions

- AI is **no longer science fiction**
  - We've been laying the scientific foundations for 70+ years
  - The biggest breakthroughs have happened in the past half decade; progress remains rapid
  - No evidence that we're on the road to artificial sentience
  - No guarantees progress won't stagnate soon
    - there's a limit to how much money can be spent on training a model
    - we're running out of training data

- Existing technologies will have big impact
  - **Mobile phones** changed society profoundly; AI will be bigger
  - **Living standards** will rise; effectively, we'll all be richer
  - Widespread unemployment not around the corner, but AI will **automate many tasks** that now need human expertise

- **Government** (and voters) will shape the agenda
  - funding; adopting technologies; **regulation**

So How Did You Make The Pictures?

- Generated by DALL-E 3
  - https://www.bing.com/images/create

- Prompt:
  - "A wide angle shot of cute, happy animals gathered around the yellow letters "AI". The animals are all different colors, and each is very detailed and has a distinct personality. The background of the image is completely white. The style is "craft clay": photorealistic 3d picture like in a claymation movie. All of the animals are clearly made of smooth, shiny clay."