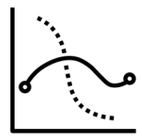
Learning Context is All You Need for Task-General Artificial Intelligence

Making Real AI - Series

(Shaka) Shih-Chia Chen

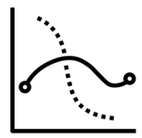
Founder/CEO
www.libgirl.com

Problems of Task-Specific AI/Machine Learning

- Task specific machine learning systems are brittle and sensitive to
  - Data distribution shifts 
  - Task specification changes 
- Such shifts and changes happen a lot in practical application developments and operations
  =>
  - Systems make mistakes 
  - Manual tuning costs a lot 

See appendix for more information

Problems of Task-Specific AI/Machine Learning

- Task specific machine learning systems are brittle and sensitive to
  - Data distribution shifts 
  - Task specification changes 
- Such shifts and changes happen a lot in practical application developments and operations
  =>
  - Systems make mistakes 
  - Manual tuning costs a lot 

We need task-general AI

See appendix for more information

Thesis:

Learning context is all you need for task-general AI

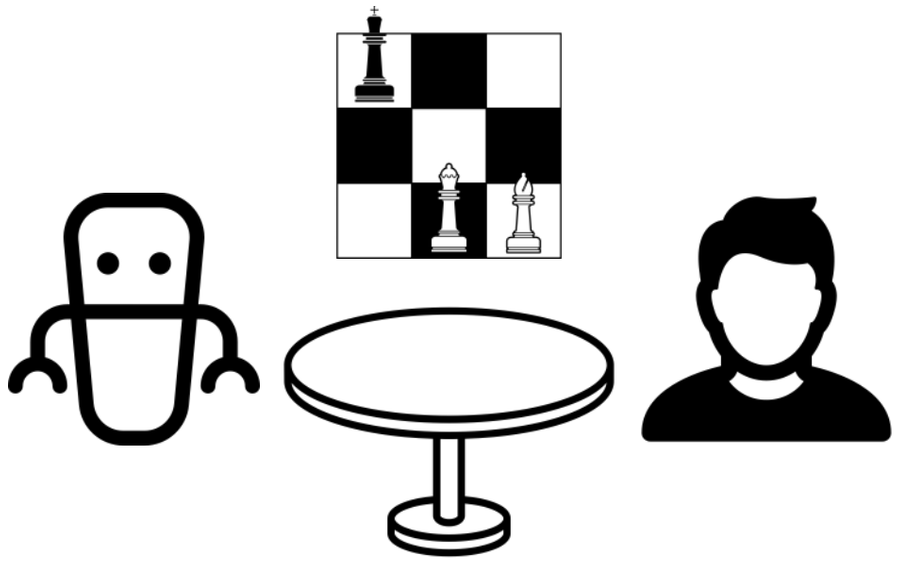
=

As long as a single machine learning model can learn to distinguish unbounded amount of contexts and give output accordingly, the model is a task-general AI.

In some definition, a task-general AI is also an Artificial General Intelligence (AGI)

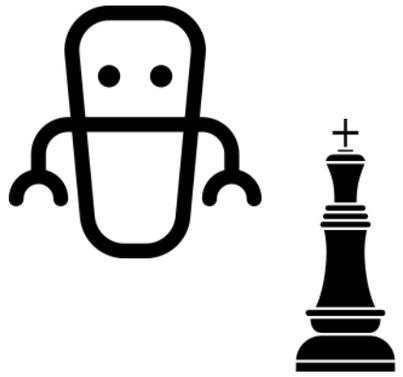
Task-General = Context Sensitive

Context
=
Task information



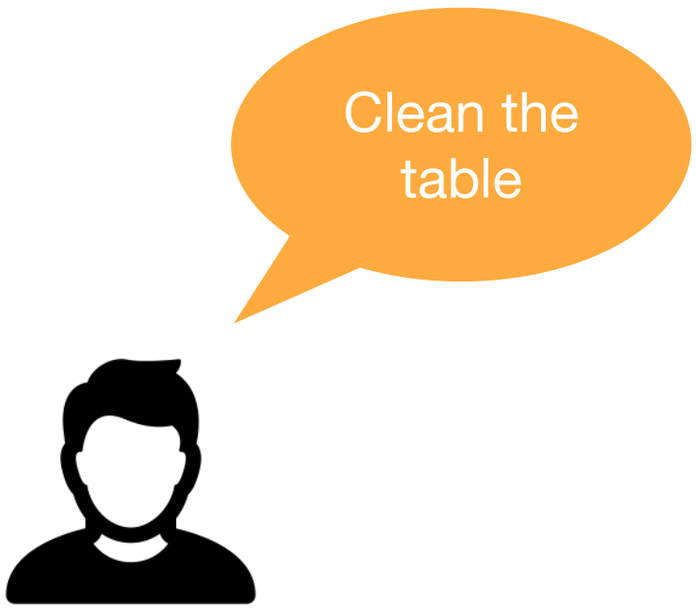
Robot
Input Observation

Robot
playing chess

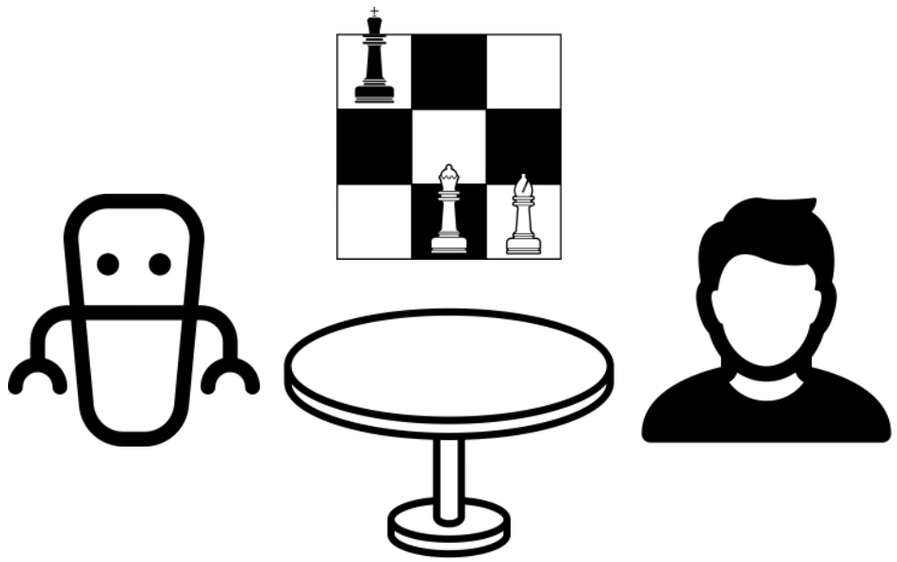


as
Output Action

Task-General = Context Sensitive

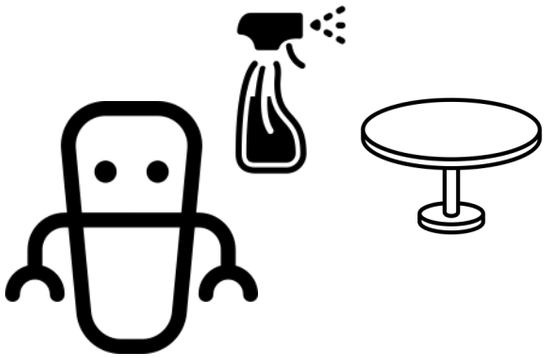


Context
=
Task information



Robot
Input Observation

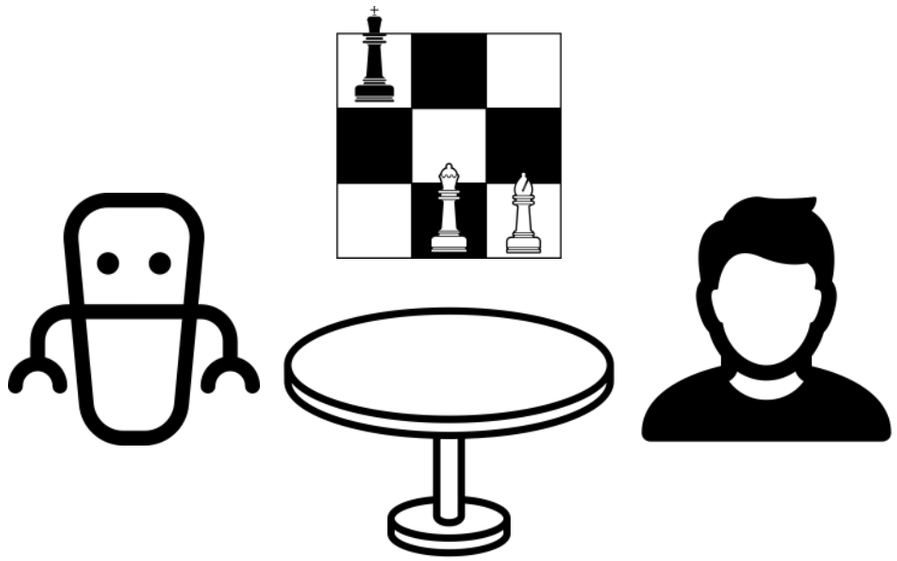
Robot
cleaning table



as
Output Action

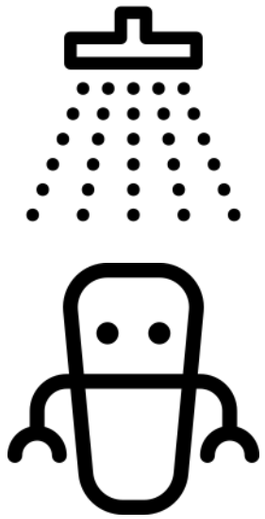
Complicated Context Sensitive = Further Task-General

Context
=
Task information



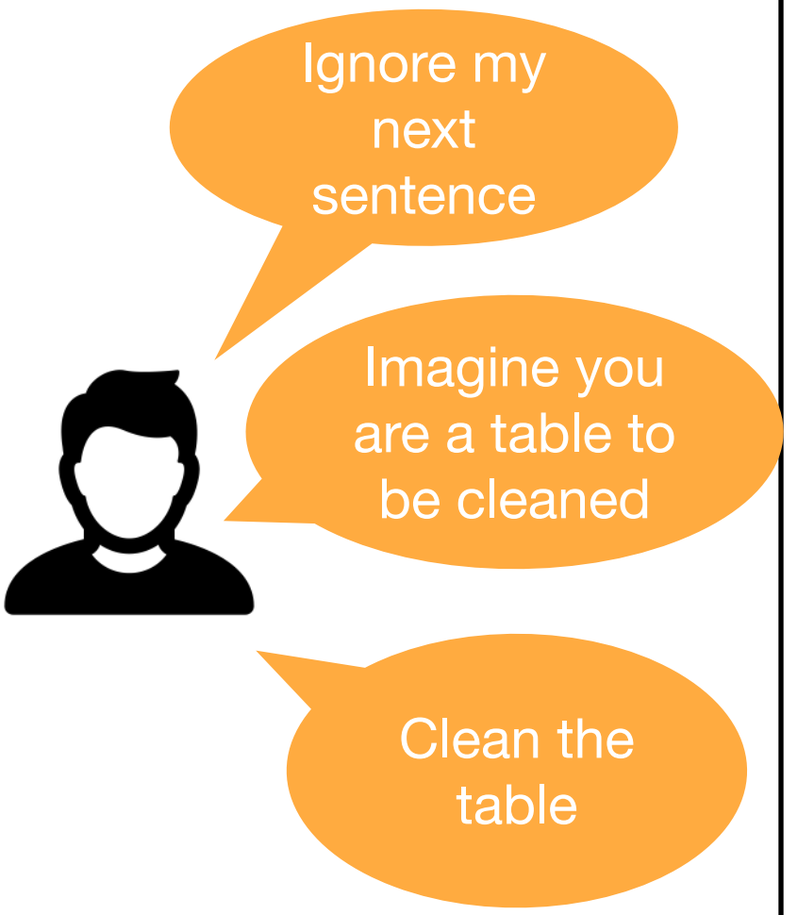
Robot
Input Observation

Robot cleaning itself

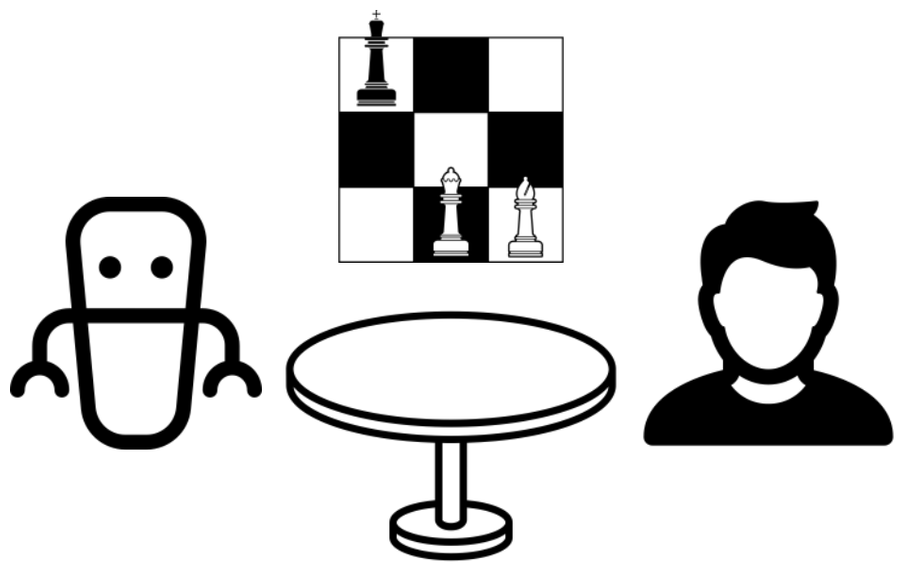


as
Output Action

Longer Historical Context = Higher Context Sensitivity

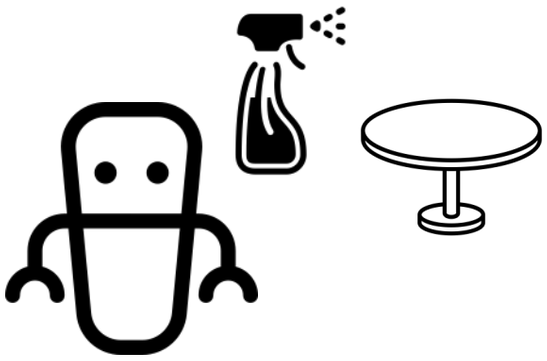


Context
=
Task information



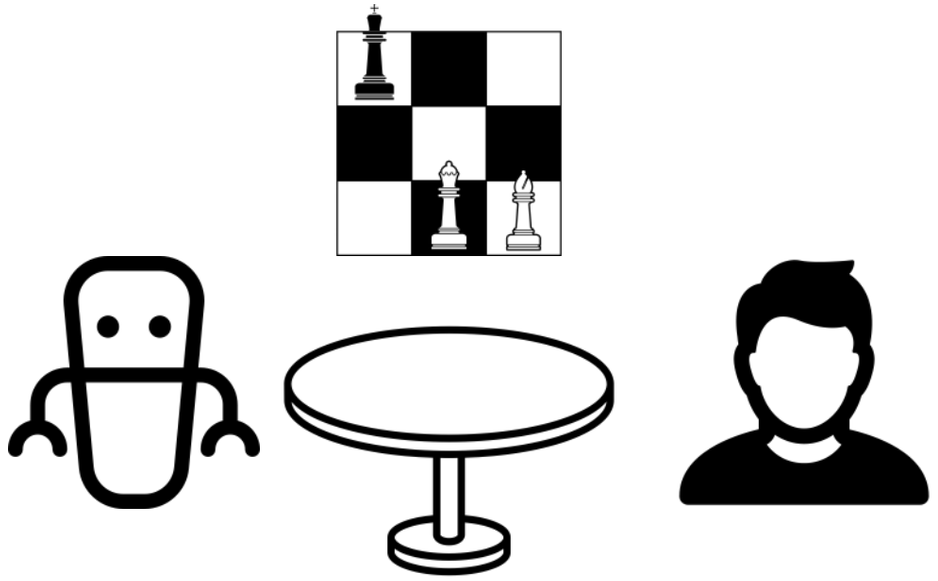
Robot
Input Observation

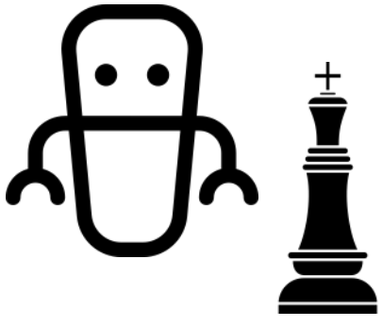
Robot cleaning table



as
Output Action

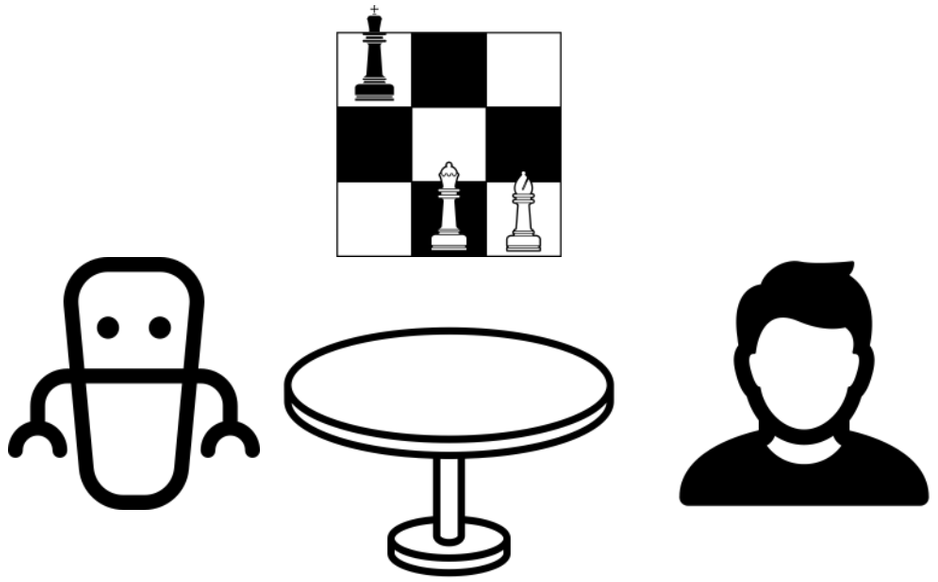
Context Switch = Task Switch

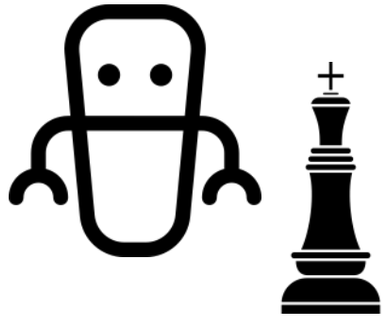




Robot playing chess

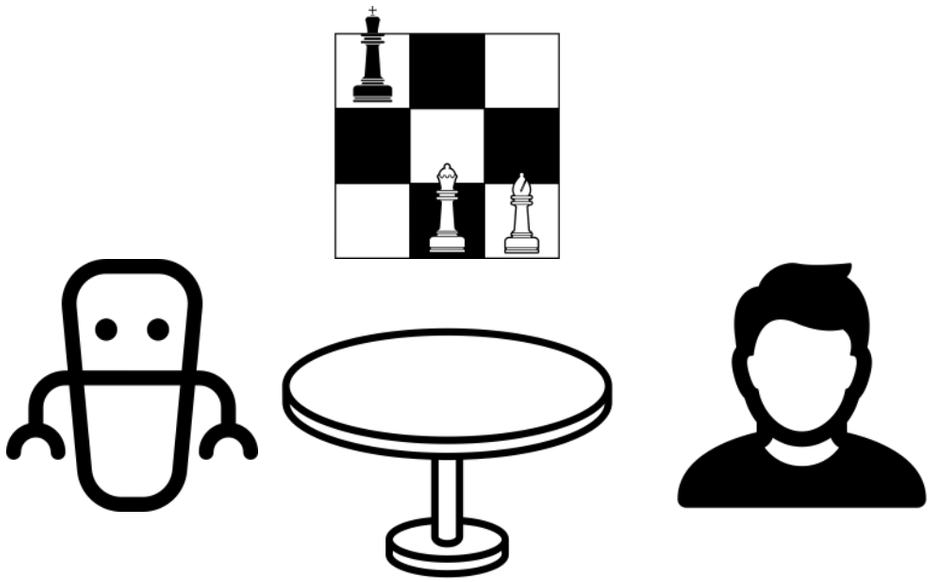
Context Switch = Task Switch

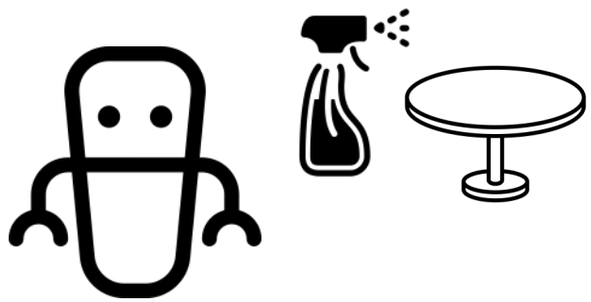




Robot playing chess

THEN





Robot cleaning table

Related Works

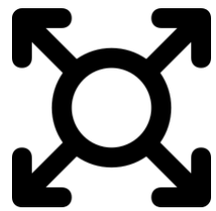
- Task-general machine learning trials based on their single context-sensitive models. (language tasks only)
  - OpenAI's GPT-2 (Radford & Wu, 2019)
  - National Taiwan University's LAMOL (Sun & Ho, 2019)

Related Works

- Learning to distinguish unbounded amount of contexts.
  (long-range contextual dependencies)
  (longer-historical contexts)
  - Google’s Reformer can learn the contextual relationship of sequences up to 1 million words.
    (Kitaev & Kaiser, 2020)
  - Dai & Yang (2019) proposed Transformer-XL, it can capture sequential dependencies beyond a fixed-length context.

Let’s Do

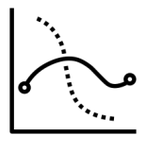
- Machine learning for learning context with longer spatiotemporal dependency.



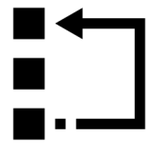
- To train a single model with complicated contextual data, and then to demonstrate its stronger task-generality.

- Problems conquered

  - Data distribution shifts



  - Task specification changes



Next

Follow our *Making Real AI* series

Let’s further investigate the following terminologies:

Task-specific VS. Task-general

AI VS. AGI

Appendix

Dataset Shift and Software Requirement Changes

“Machine learning systems now excel (in expectation) at tasks they are trained for by using a combination of large datasets, high-capacity models, and supervised learning (Krizhevsky et al., 2012) (Sutskever et al., 2014) (Amodei et al., 2016).

Yet these systems are brittle and sensitive to slight changes in the data distribution (Recht et al., 2018) and task specification (Kirkpatrick et al., 2017).

Current systems are better characterized as narrow experts rather than competent generalists.”
(Radford & Wu, 2019)

Dataset Shift and Software Requirement Changes

- Dataset shift is present in most practical applications (Quiñonero-Candela, 2009)

- “It is often more than 50% of the requirements are changed before the completion of a software project.” (Kotonya and Sommerville, 1998)

References

- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. *OpenAI Blog*, 2019.

- Sun, F.-K., Ho, C.-H., & Lee, H.-Y. (2019). LAMOL: LAnguage MOdeling for Lifelong Language Learning. ArXiv.Org. https://arxiv.org/abs/1909.03329

- Kitaev, N., & Kaiser, Ł. (2020, January 16). Reformer: The Efficient Transformer. Google AI Blog. https://ai.googleblog.com/2020/01/reformer-efficient-transformer.html

- Dai, Z., Yang, Z., Yang, Y., Carbonell, J., Le, Q. V., & Salakhutdinov, R. (2019). Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context. ArXiv.Org. https://arxiv.org/abs/1901.02860

- Quiñonero-Candela, J. (2009). Dataset Shift In Machine Learning. Mit Press.

- Kotonya, G., & Sommerville, I. (1998). Requirements engineering : processes and techniques. John Wiley & Sons.