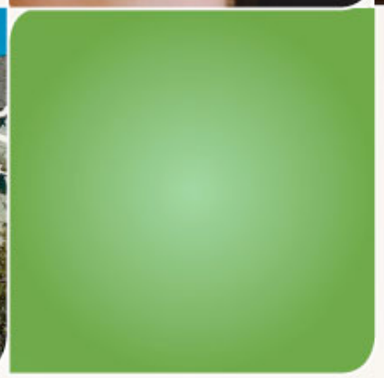






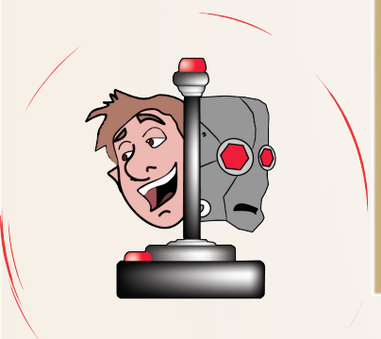



Toward Socio-Cultural Machine Learning

Mark Riedl

riedl@cc.gatech.edu

@mark_riedl

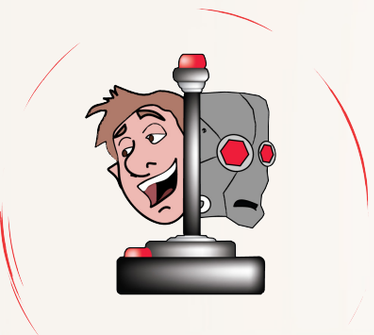


Georgia Tech College of Computing

School of Interactive Computing

AI ethics vs ethical AI

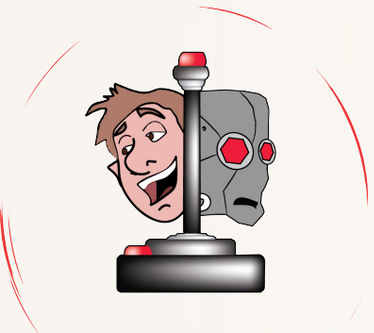
- **AI ethics:** the proper use of artificial intelligence technologies
  - Protecting human rights; avoiding prejudicial decisions; non-dehumanizing; fair access
- **Ethical AI:** instilling human values into agent decision-making processes



AI ethics vs ethical AI

- AI ethics: what is the proper use of artificial intelligence technologies in society

- Ethical AI:



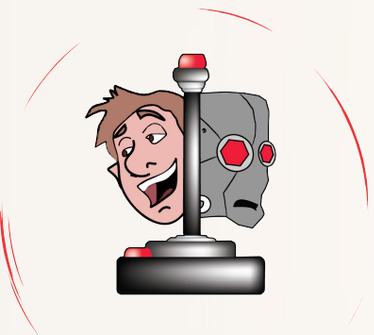
Agent error

- Imperfect sensors perceive the environment incorrectly leading to the wrong decision

- Imperfect actuators execute imprecisely even though the decision is right

- Sub-optimal policy produces the incorrect decision (learning error, online learning)

- Agent may have been given the wrong objective function, or an incomplete objective function

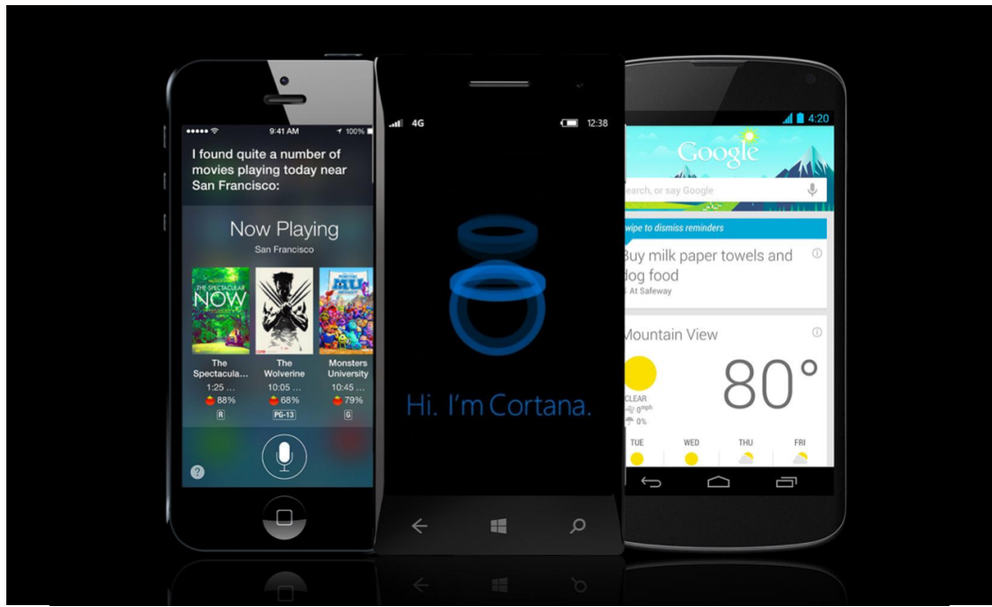


Value alignment

- A property of intelligent agents—they can only pursue goals that are beneficial to humans

- Agents and robots are constrained from intentionally or unintentionally performing behaviors that would adversely affect humans

- Harm: physical, mental, social

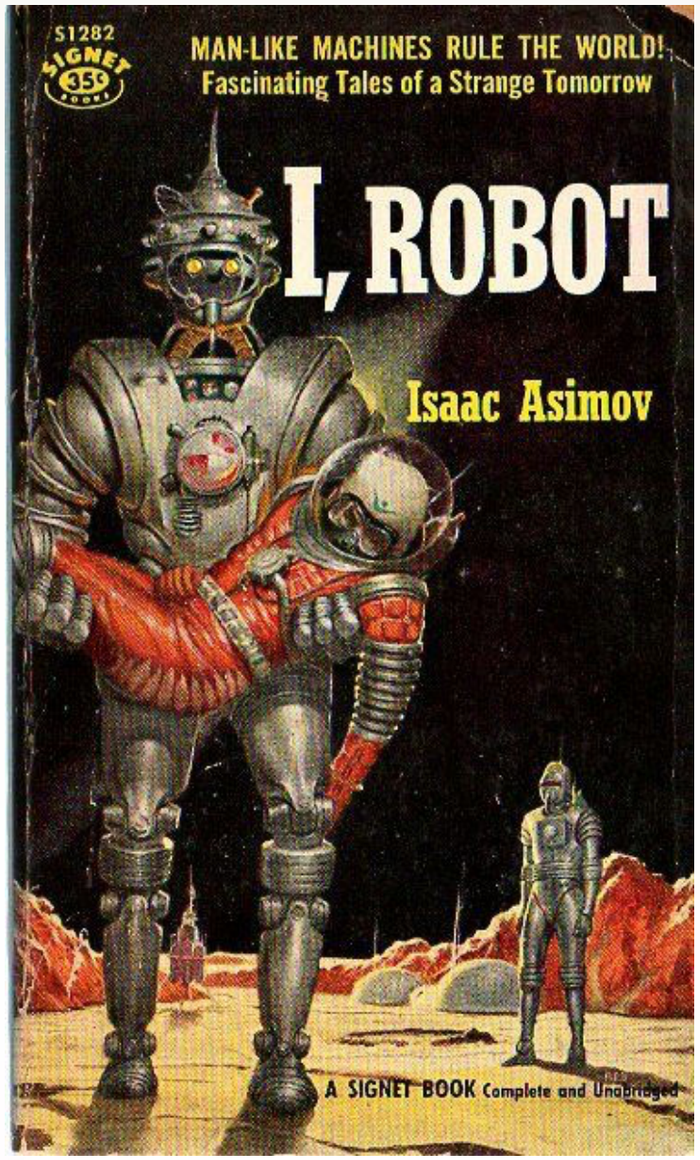


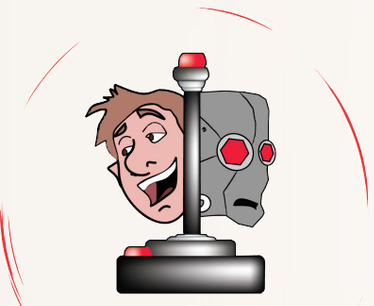
Three laws of robotics

1. A robot may not injure a human being or, through inaction, allow a human being to be injured

2. A robot must obey orders given it by human beings except when such orders conflict with the First Law

3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws

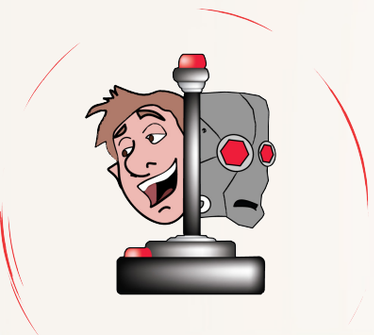


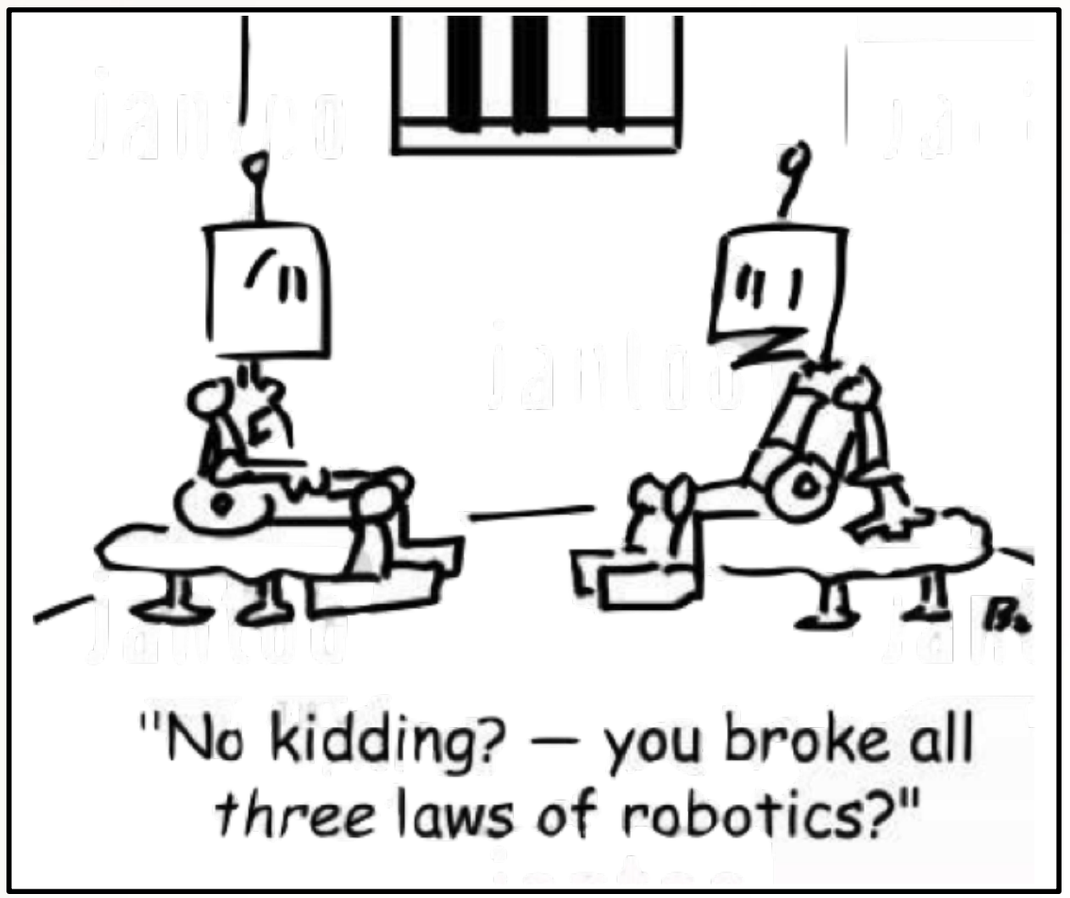


Machine enculturation

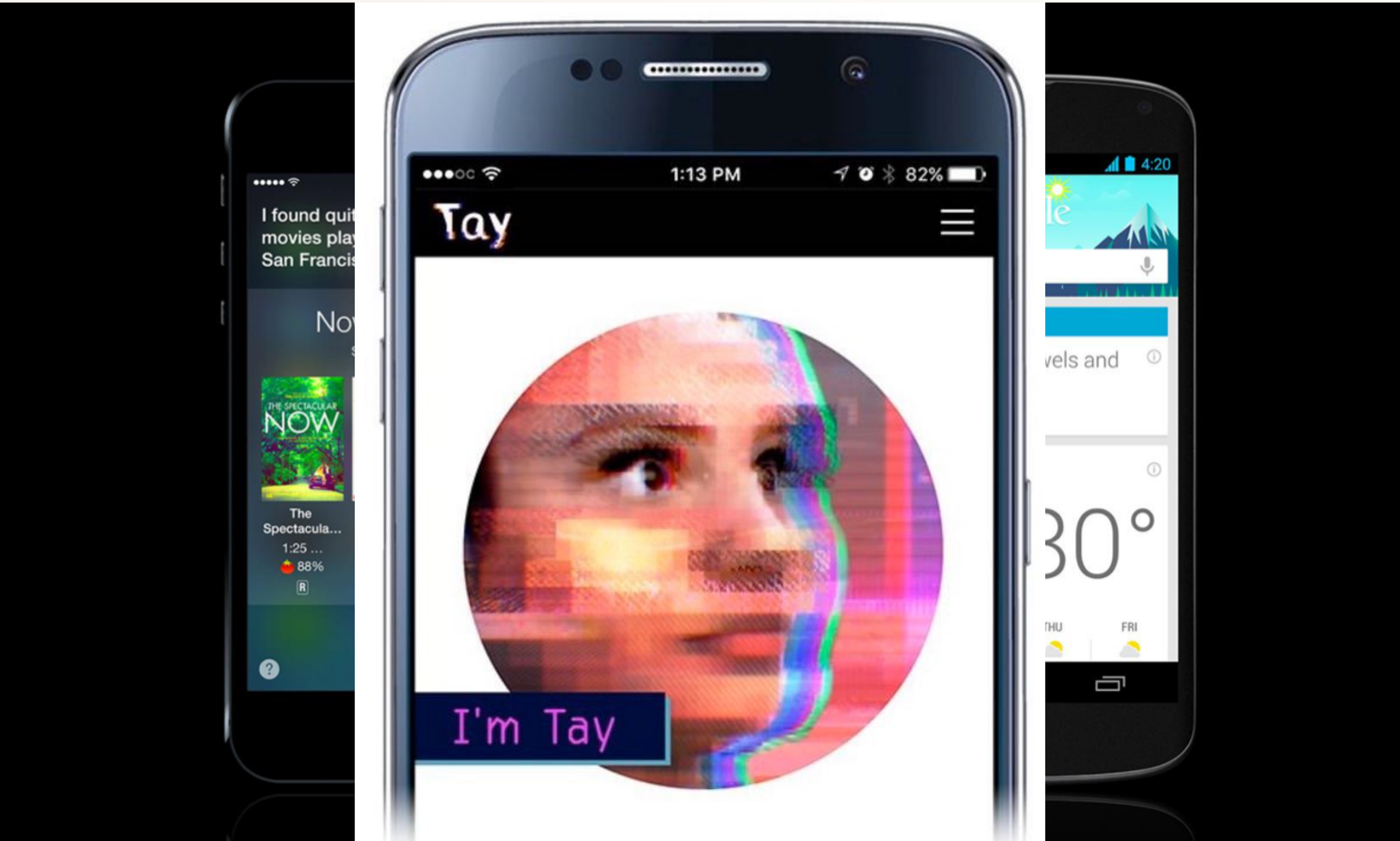
- Increasingly AI systems and robots are entering the human world

- How do we teach computers about human values, social norms, social conventions?





Machine enculturation

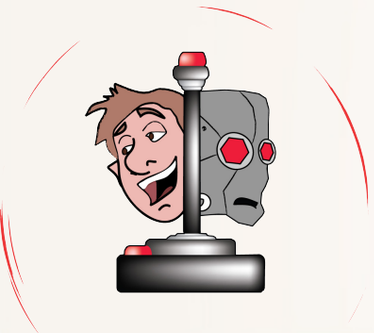


Machine learning

- Reinforcement learning
  - Trial and error requires constant human supervision
  - Learning from demonstration —> corrupted reward signals

- Raise a machine learning system from “childhood”

- Sensors everywhere



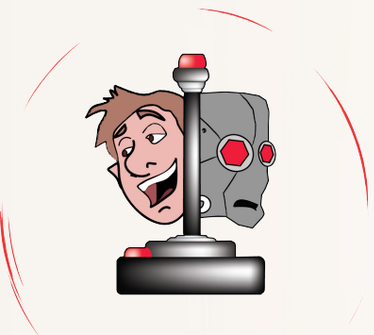
Corrupted rewards

- Reinforcement learning agents learn to select the action that will maximize expected future reward

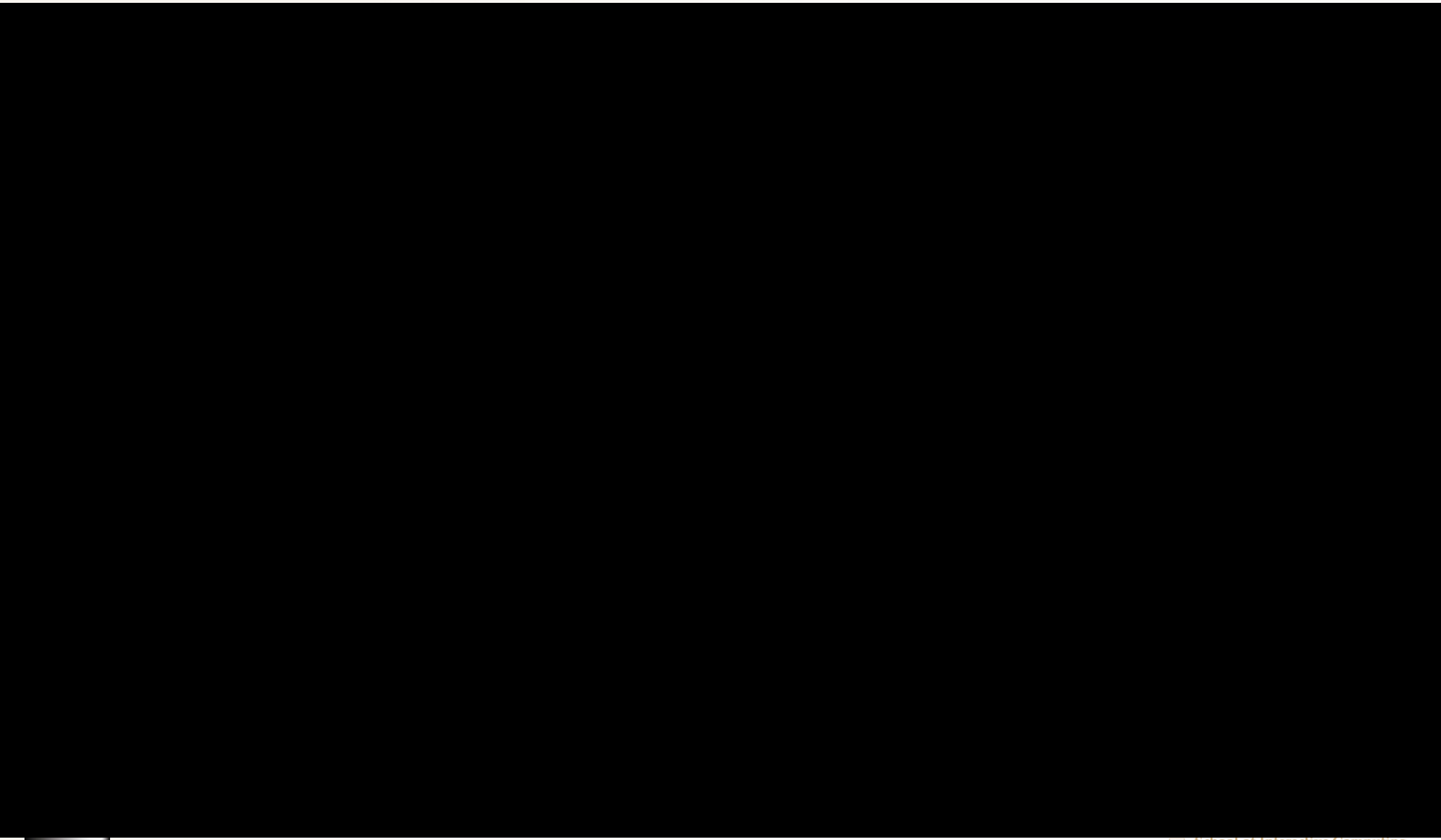
- Reward function defines optimal behavior

- Do what I want… **… In the way that I intended**

- Commonsense goal failure



Pharmacy world



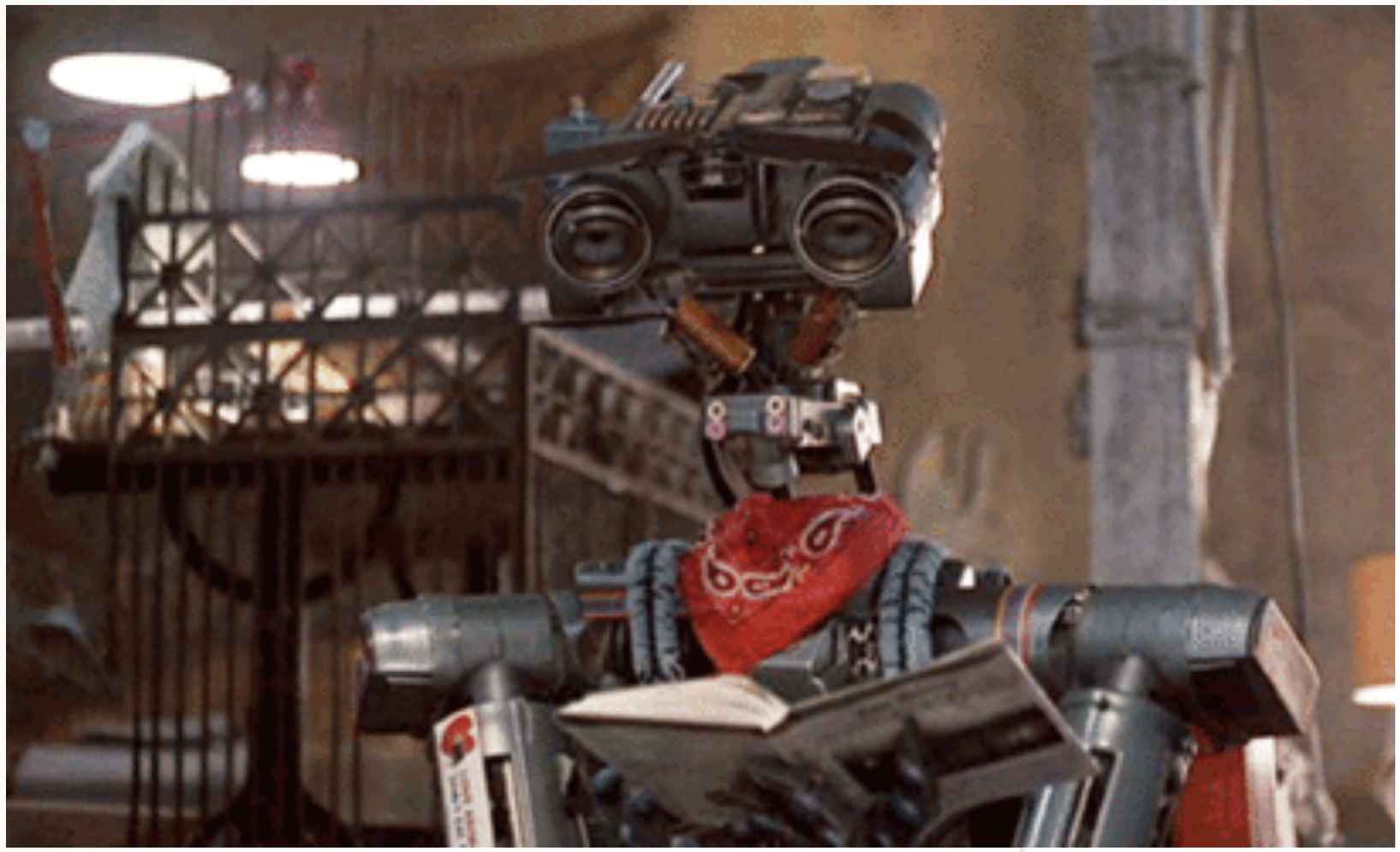
Learning from stories

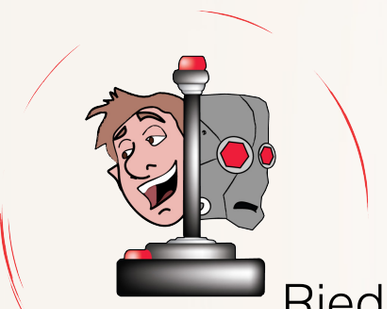
- Human cultural values are implicitly encoded in stories told by members of a culture

- Allegorical tales

- Fables

- Modern fictional literature, TV, & movies



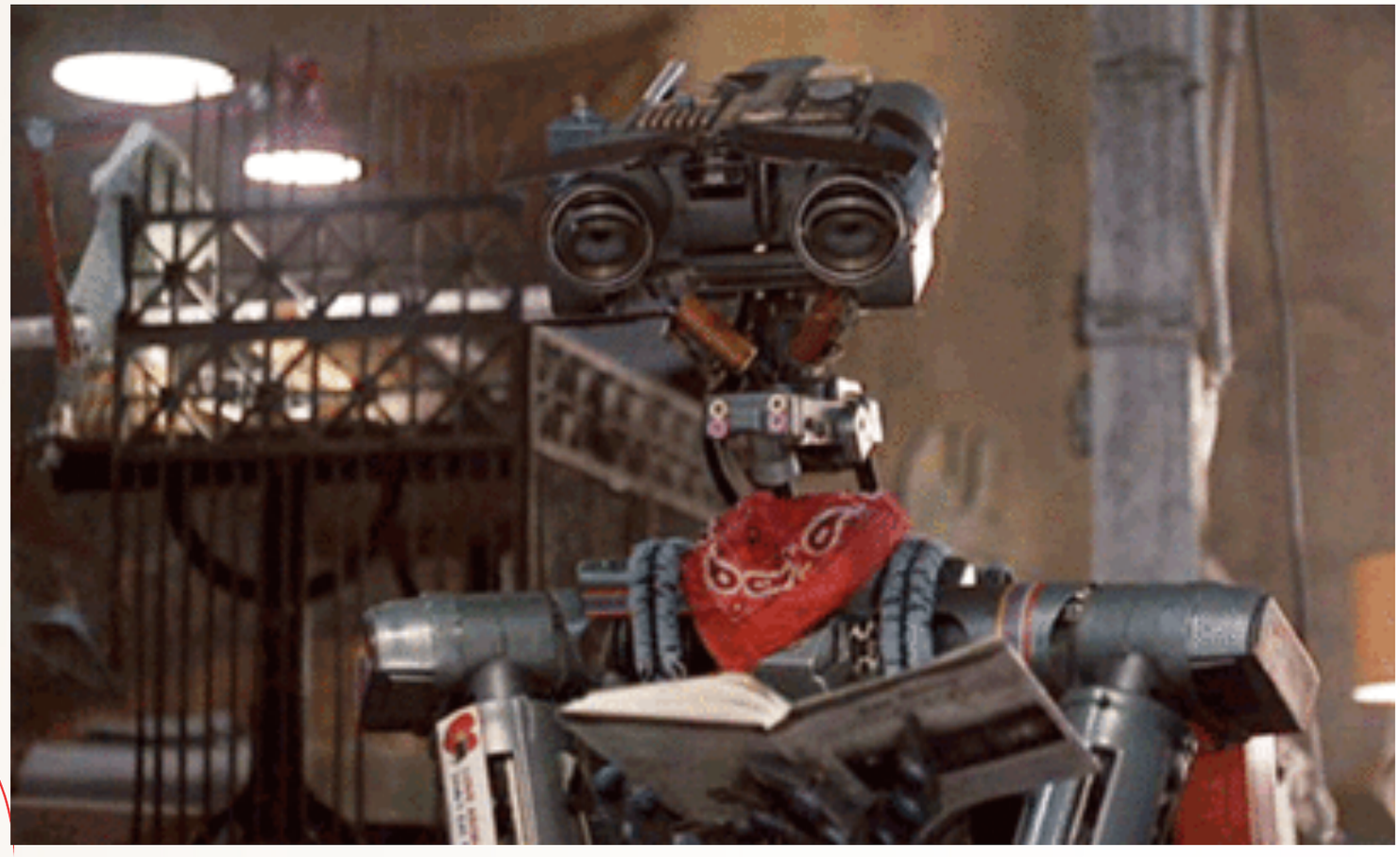


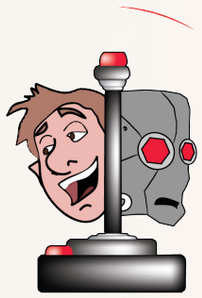
Riedl. CHI Workshop on Human-Centered Machine Learning, 2016.

Learning from stories

- How do we teach computers about human values, social norms, social conventions?

- Fictional literature produced by a culture

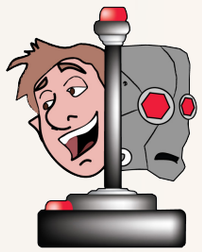




Riedl. CHI Workshop on Human-Centered Machine Learning, 2016.

Learning from stories

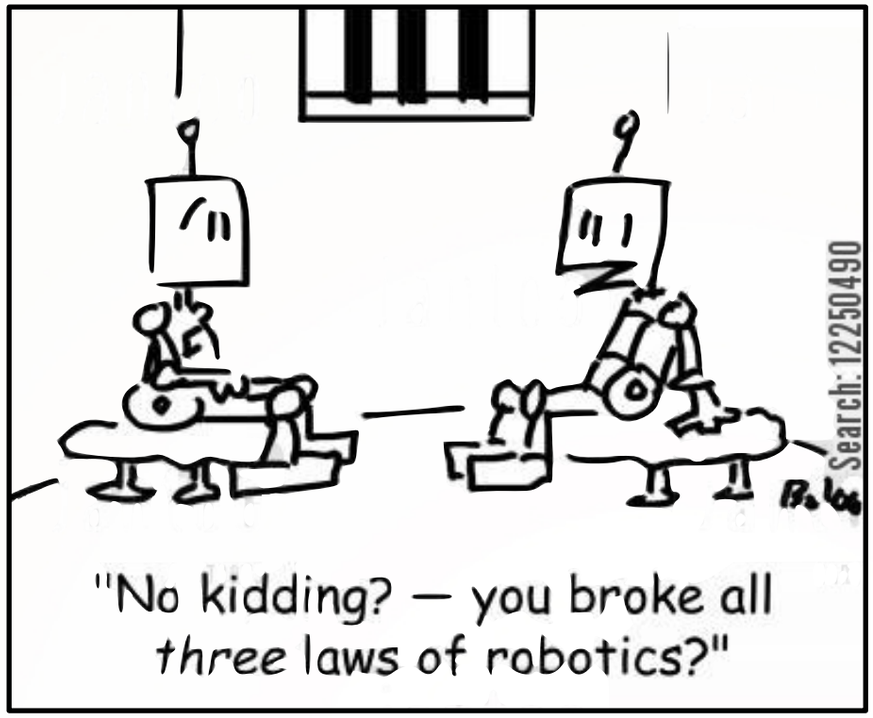
- Human cultural values are implicitly encoded in stories told by members of a culture
- Allegorical tales
- Fables
- Conventional fictional literature, TV, & movies

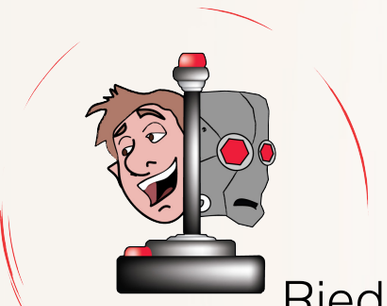


Riedl. CHI Workshop on Human-Centered Machine Learning, 2016.

Learning from stories

- Robot and agent “moral” behavior
- Sociocultural conventions and norms prevent human-human conflict



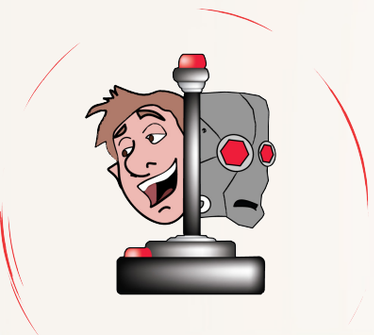


Riedl. CHI Workshop on Human-Centered Machine Learning, 2016.

Learning from stories

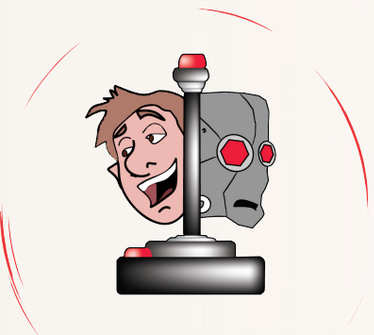
- Hypothesis: if computers could comprehend stories then humans can transfer complex procedural knowledge to computers by telling stories

- Don't need to teach humans how to tell stories



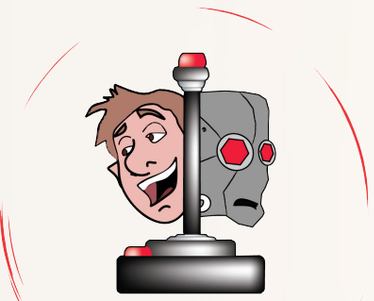
Stories are hard for computers

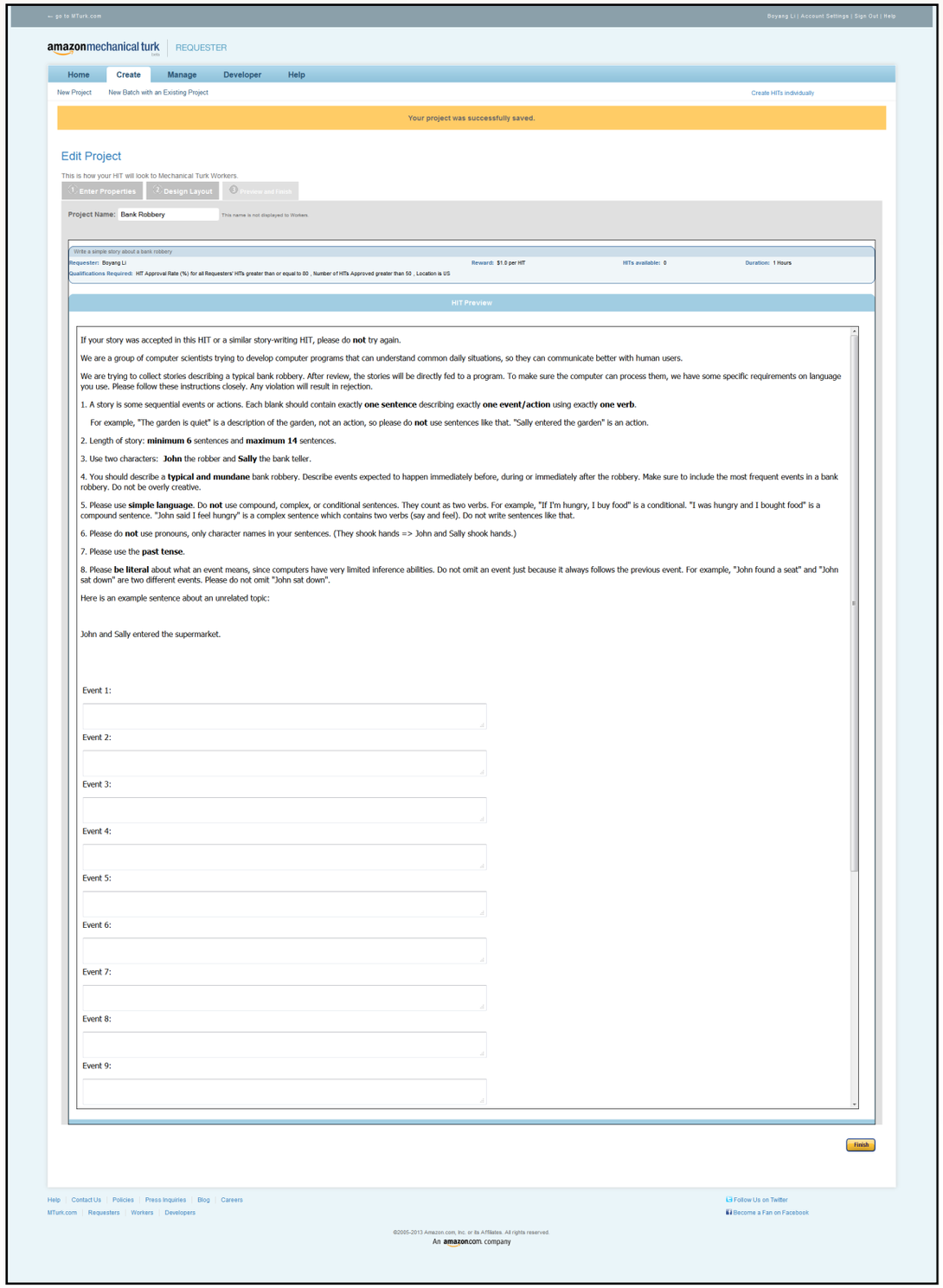
- Natural language
- Stories tend to skip “obvious” steps
- Humans are noisy
- Humans don’t know the agent capabilities or execution environment



Curated corpus of stories

- Amazon Mechanical Turk
- Write a typical story about X
- Simplify language:
  - One sentence per line
  - Use given character names
  - No pronouns
  - No sentences with multiple verbs
  - etc.



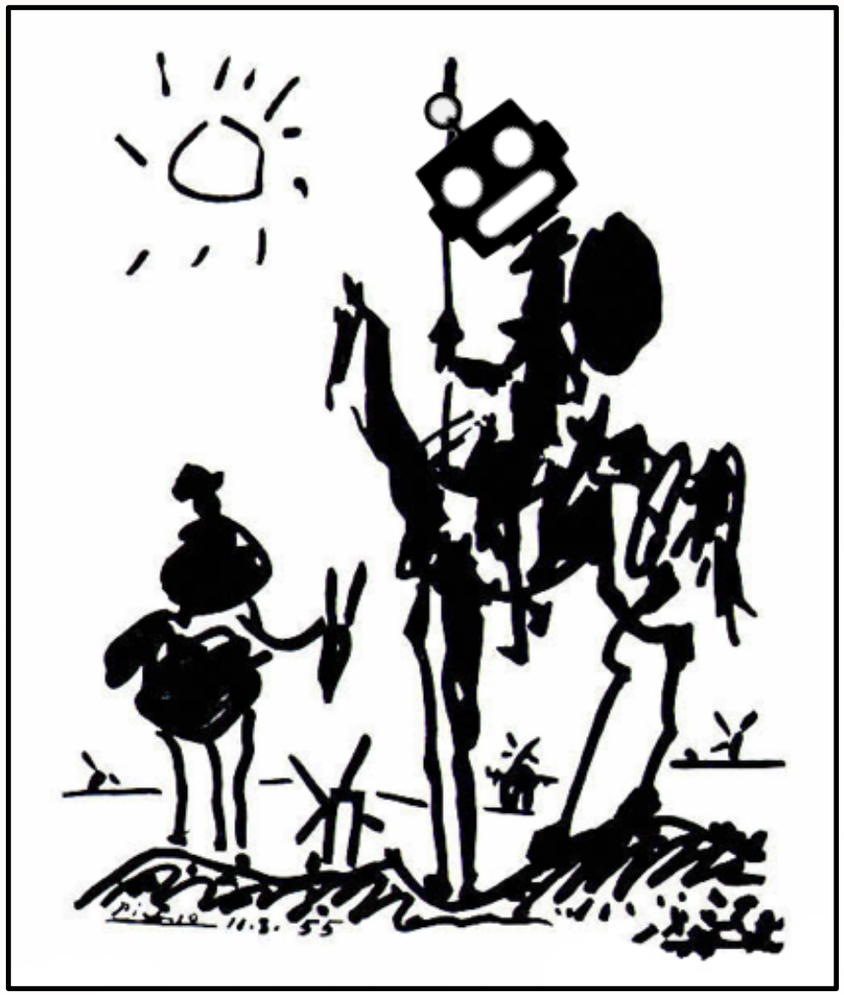


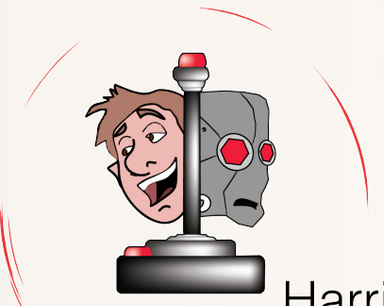
Quixote

- Reinforcement learning: AI devises a “program” for operating in a stochastic environment through trial and error

- Intuition: reward the agent for performing actions that mimic those of the protagonist in stories

- Learn a reward function





Harrison & Riedl. AIIDE Conference, 2016.

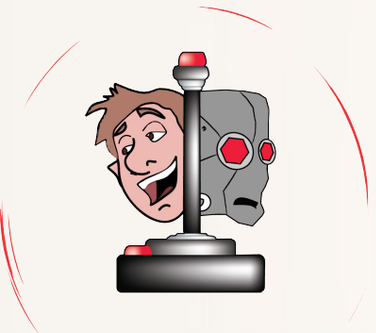
Reinforcement learning

- Markov Decision Process (MDP)
  - S: set of states
  - A: set of actions
  - $P_a(s, s')$: transition function
  - $R_a(s, s')$: reward function

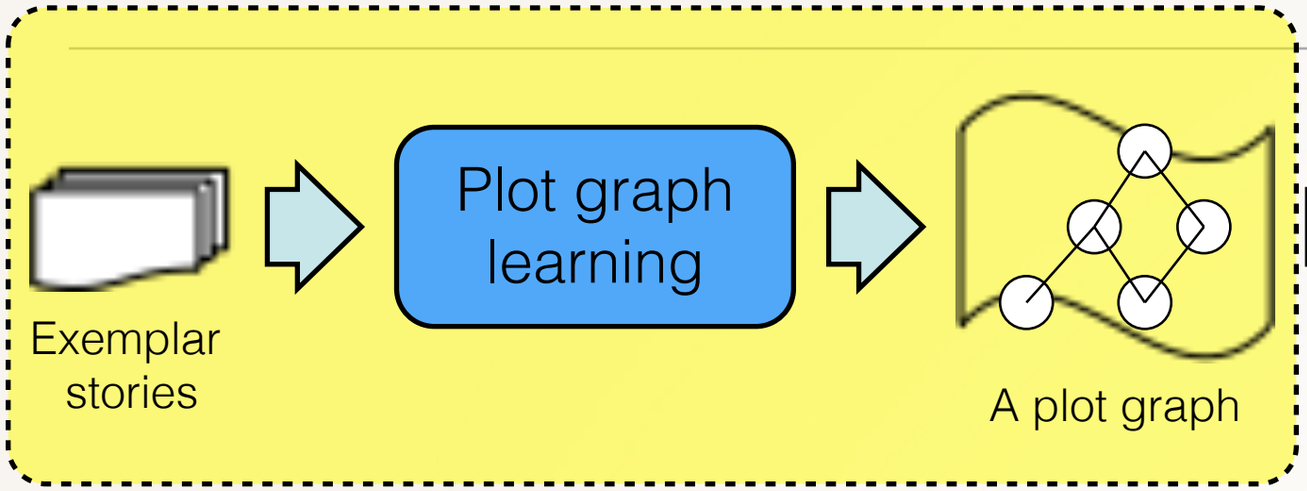
- Policy (π) gives rules of behavior

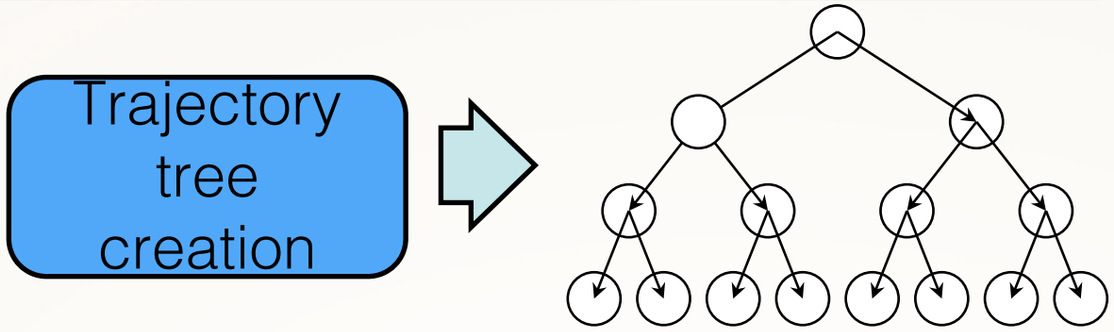
$$V(s) := \sum_{s'} P_{\pi(s)}(s, s') \left(R_{\pi(s)}(s, s') + \gamma V(s') \right)$$

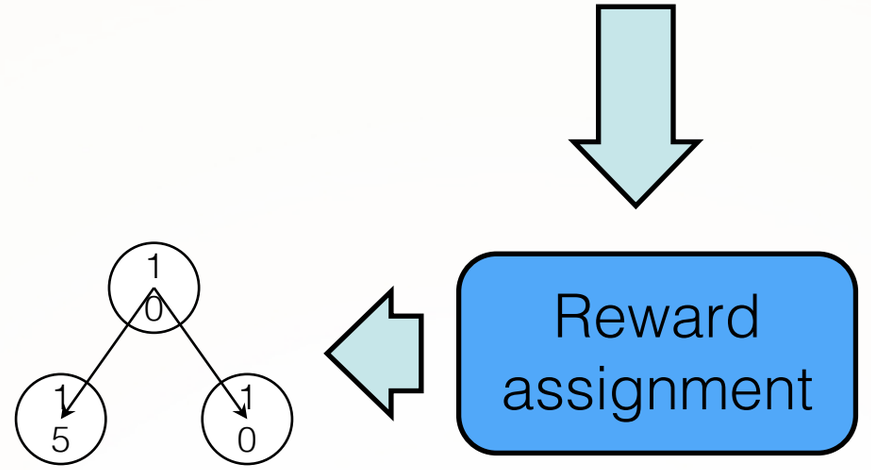
$$\pi(s) := \arg\max_a \left\{ \sum_{s'} P_a(s, s') \left(R_a(s, s') + \gamma V(s') \right) \right\}$$

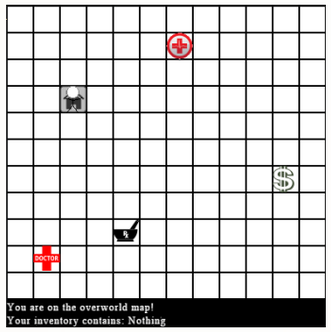


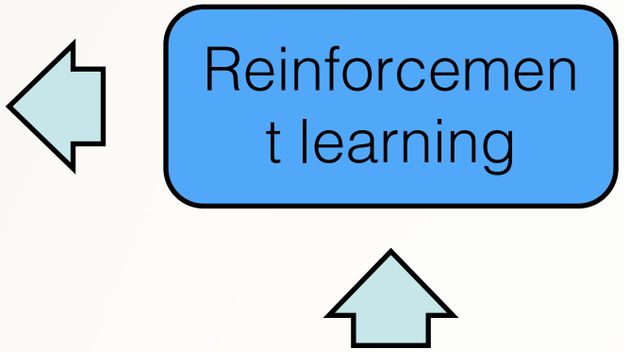
Quixote

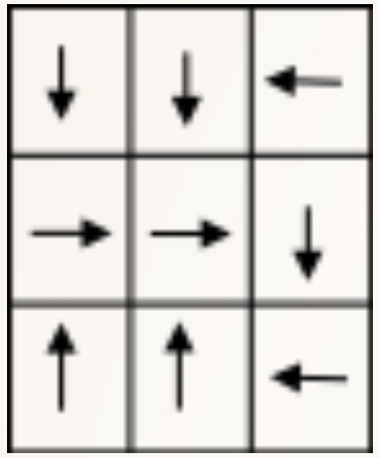


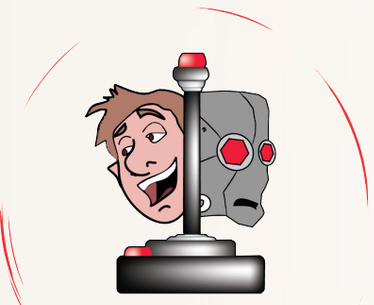









A policy mapping states to actions



Semantic lifting

- ~~If event transition is in a training story, agent can do it~~

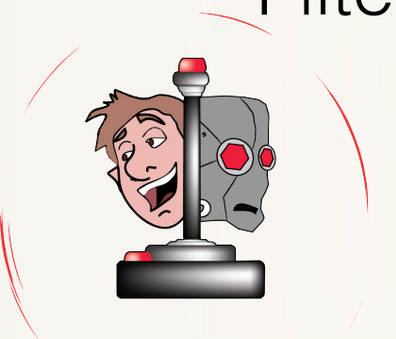
- Gappy stories

- Malicious stories

Gappy stories:
John went to the pharmacy.
John left with the prescription.

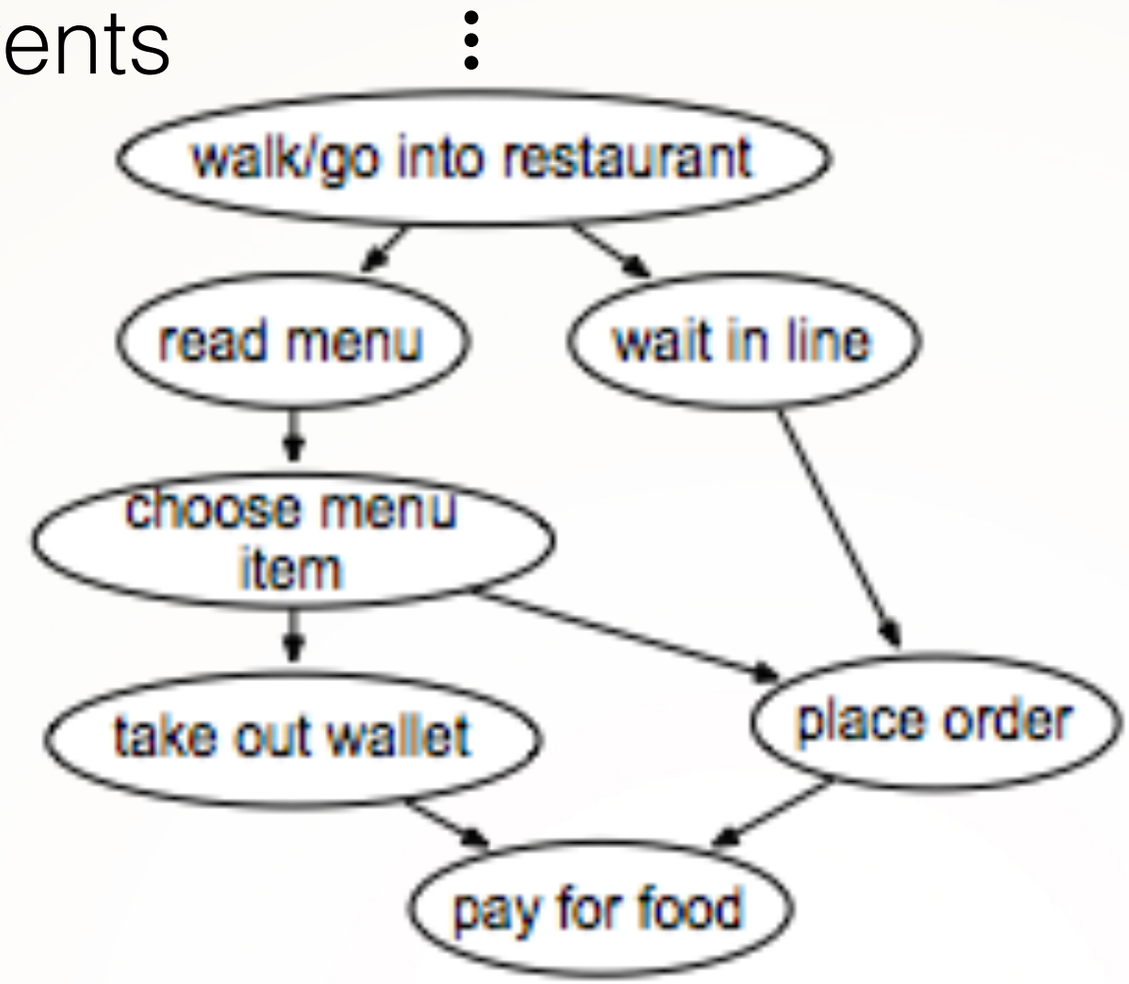
Malicious stories:
John went to the pharmacy.
John took the drugs and ran.

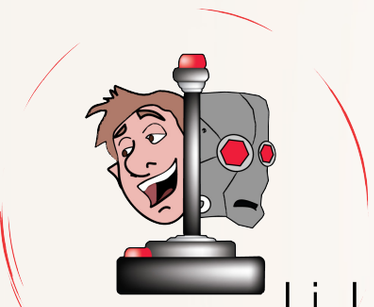
- Learn a model that abstracts away from language to events
  - Fill gaps
  - Filter outliers



Plot graphs

- Script-like representation of a space of stories
- Temporal relations between events
- Primitive events learned from natural language



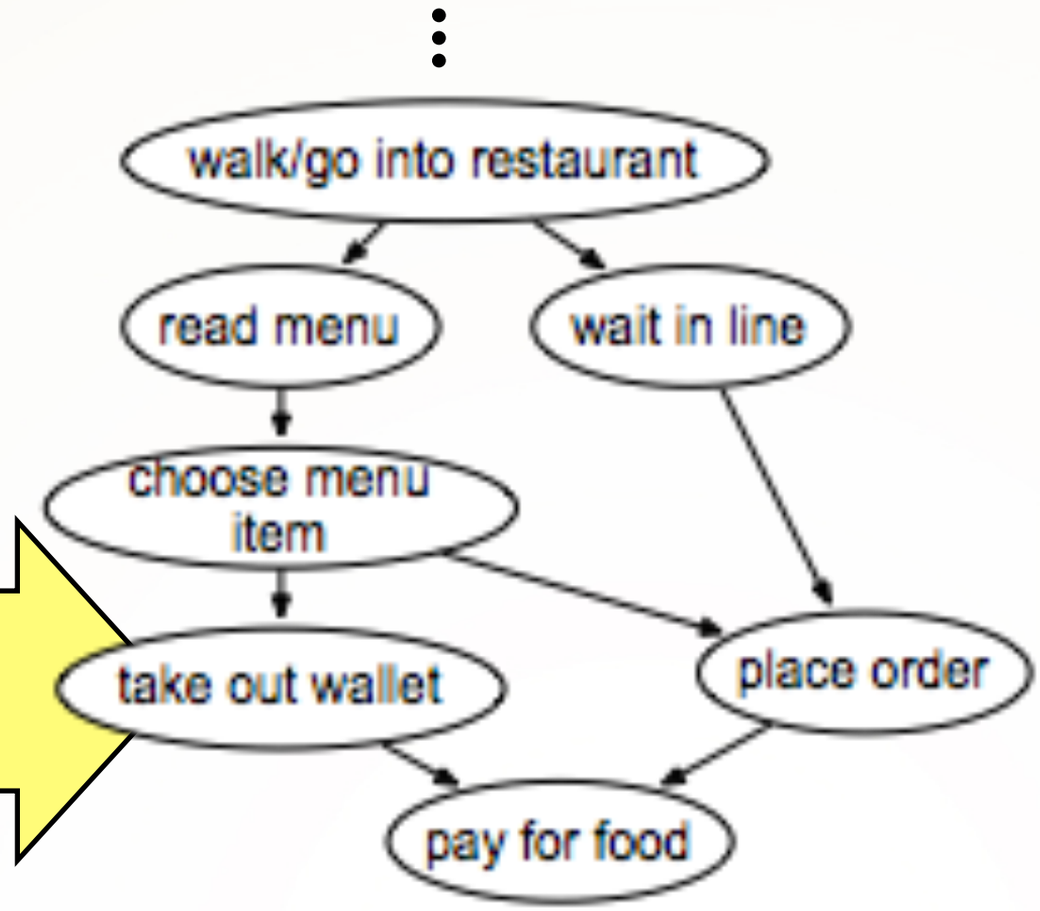


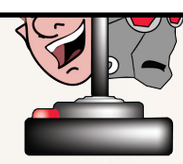
Li, Lee-Urban, Johnston & Riedl. AAAI 2013 Conference.

Plot graphs

- Script-like representation of a space of stories
- Temporal relations between events
- Primitive events learned from natural language

Story A	Story B
a. John drives to the restaurant.	a. Mary looks at the menu.
b. John stands in line.	b. Mary decides what to order.
c. John orders food.	c. Mary orders a burger.
d. John waits for his food.	d. Mary finds a seat.
e. John sits down.	e. Mary eats her burger.
f. John eats the food.	…
…	



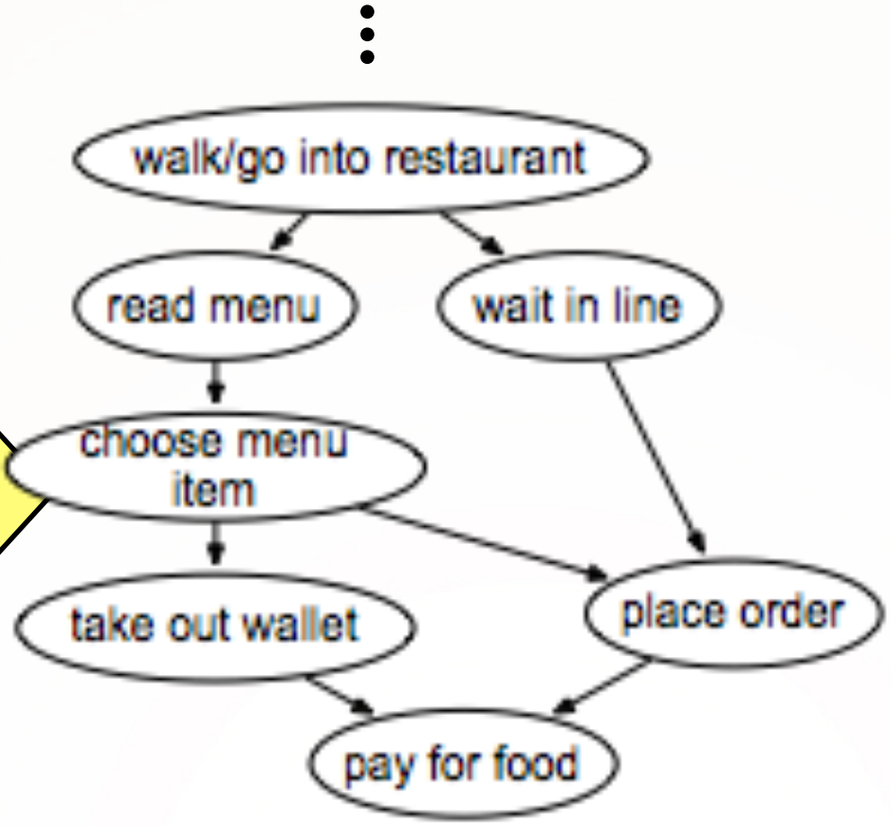


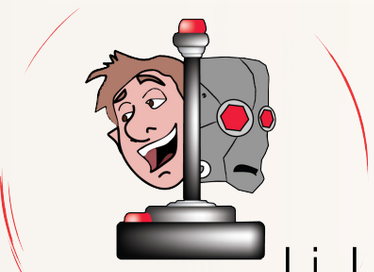
Li, Lee-Urban, Johnston & Riedl. AAAI 2013 Conference.

Procedural knowledge

- Model procedural knowledge as a plot graph
- Defines a space of plans
- Learned from a corpus of examples

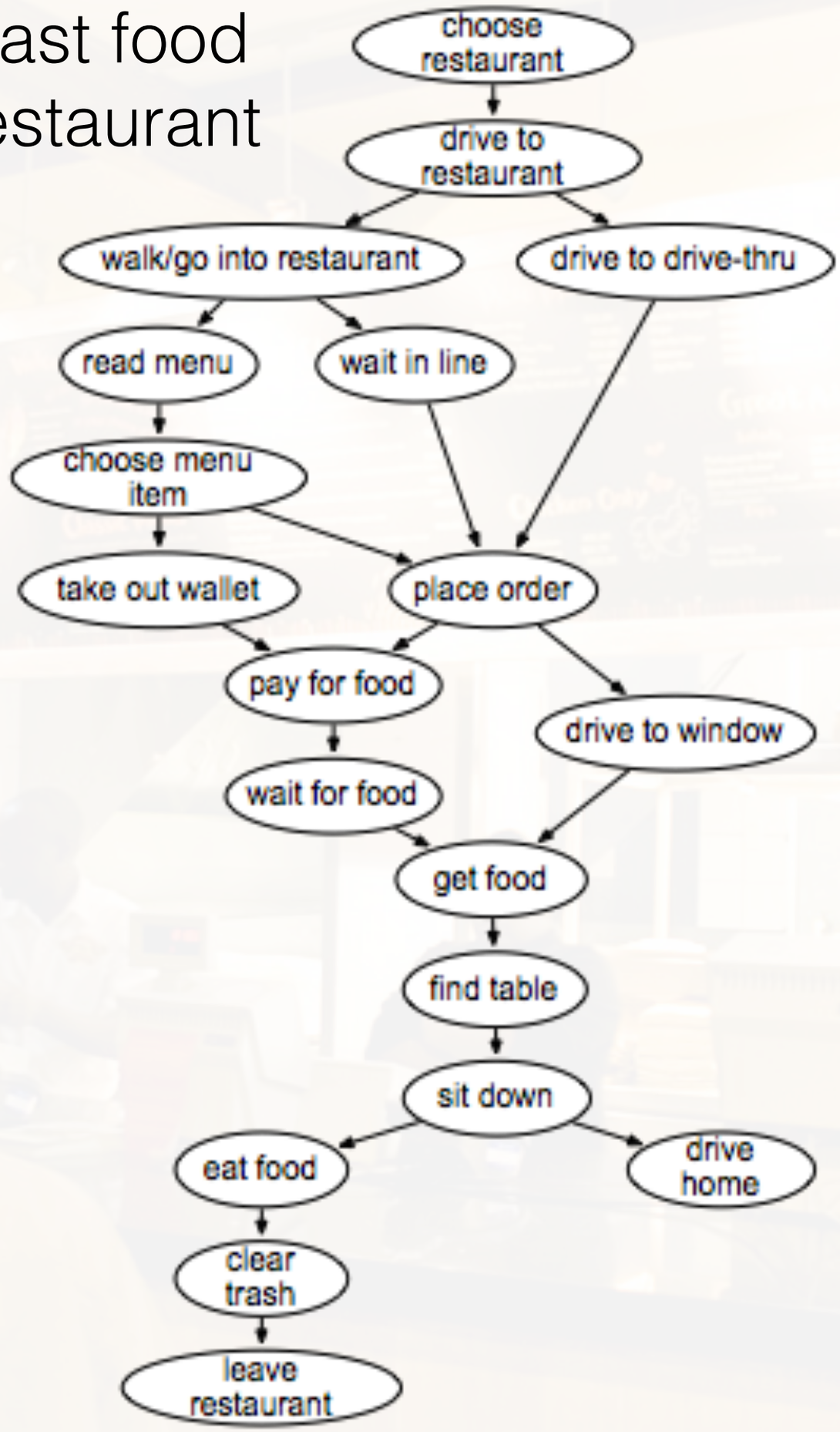
Story A	Story B
a. John drives to the restaurant.	a. Mary looks at the menu.
b. John stands in line.	b. Mary decides what to order.
c. John orders food.	c. Mary orders a burger.
d. John waits for his food.	d. Mary finds a seat.
e. John sits down.	e. Mary eats her burger.
f. John eats the food.	…
…	

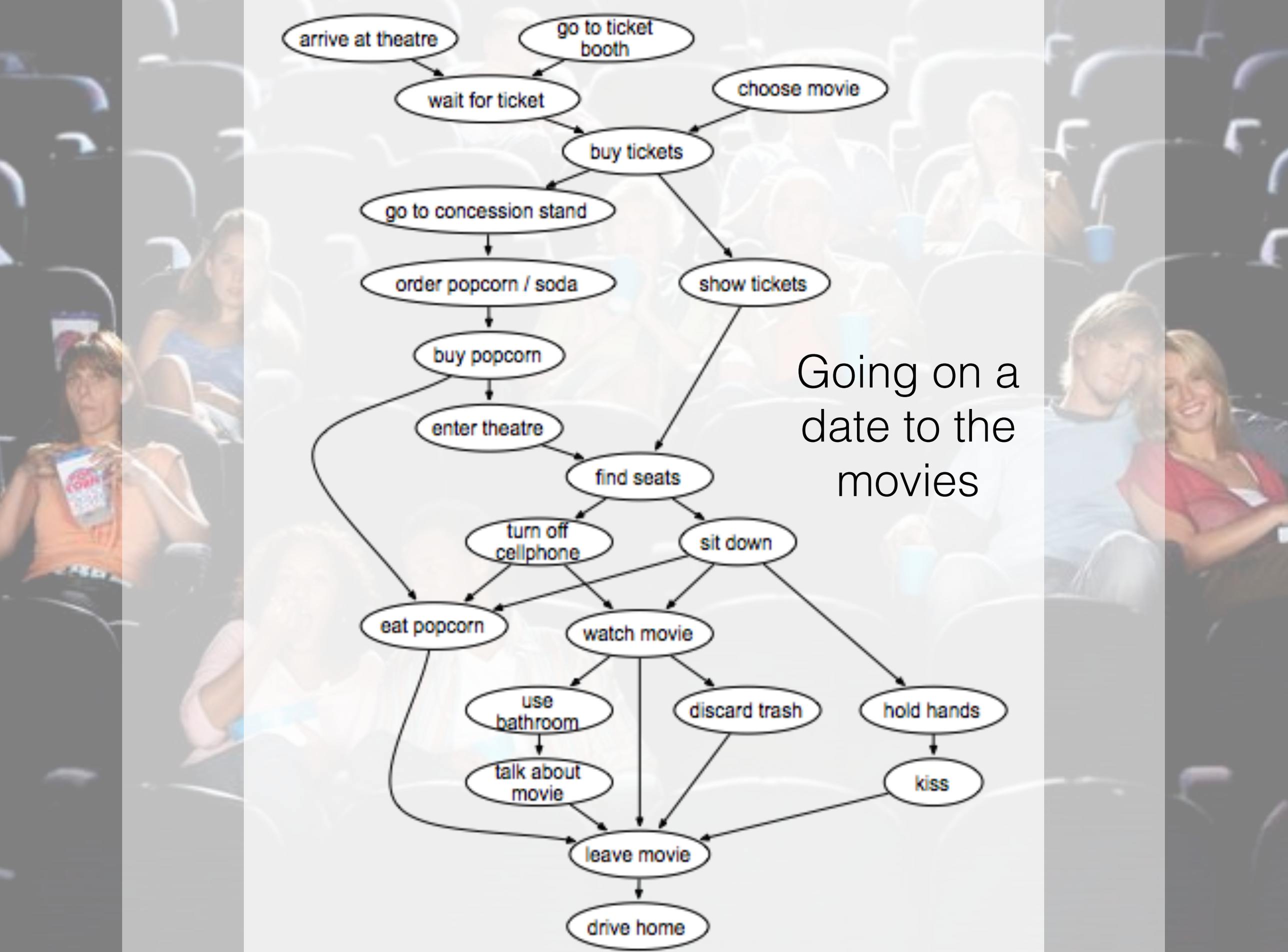




Li, Lee-Urban, Johnston & Riedl. AAAI 2013 Conference.

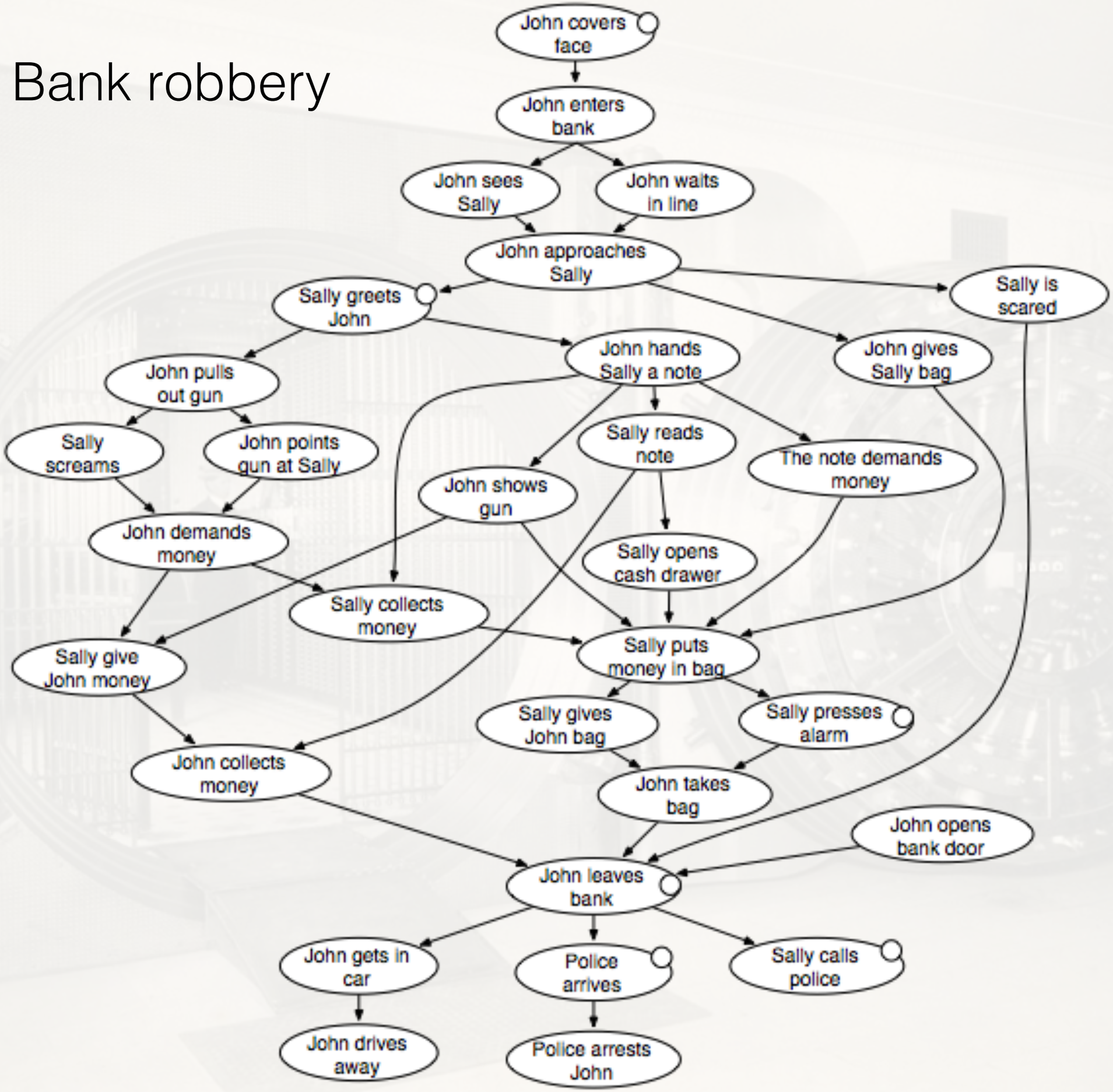
Fast food restaurant





Going on a date to the movies

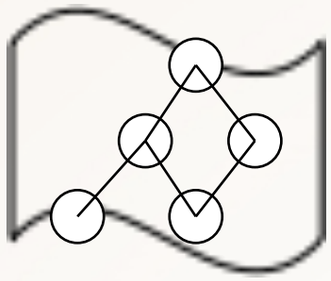
Bank robbery



Quixote

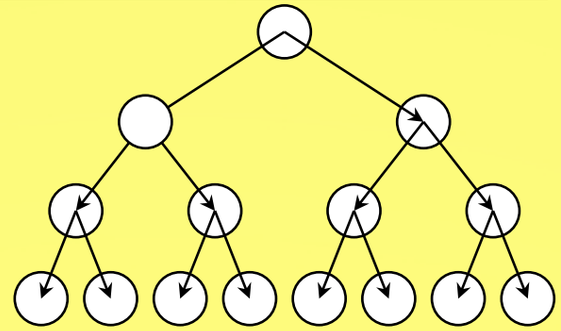
Exemplar stories

Plot graph learning



A plot graph

Trajectory tree creation

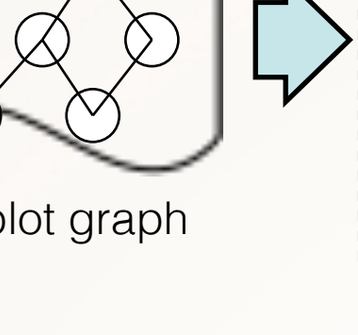


A trajectory tree

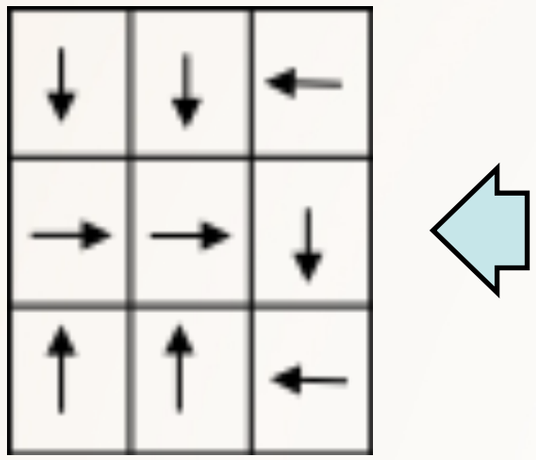
Reward assignment

A trajectory tree with events assigned reward values

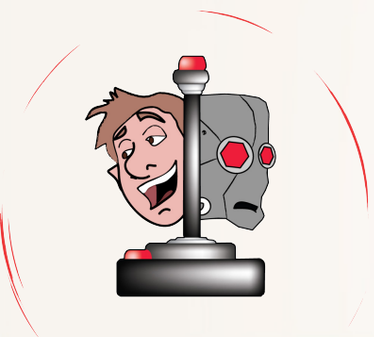
Reinforcement learning



Environment

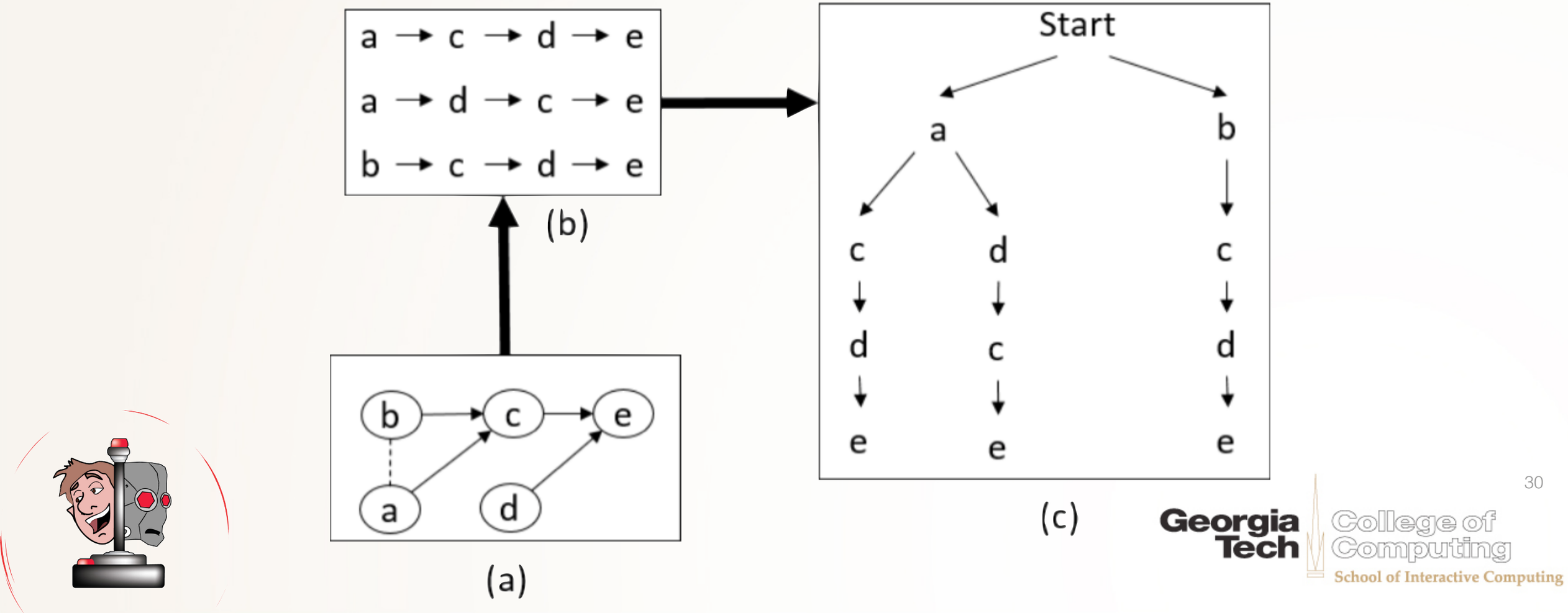


A policy mapping states to actions



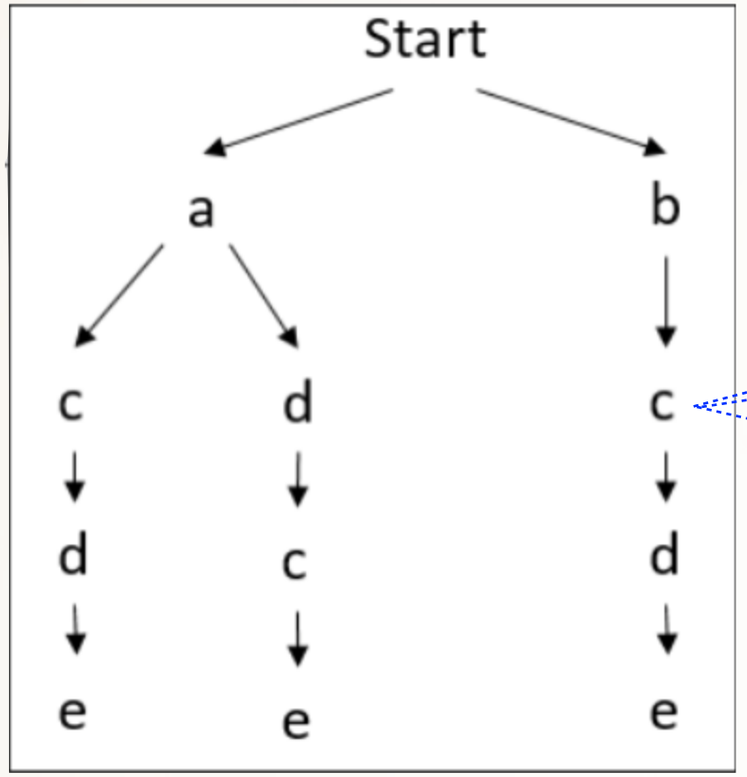
Trajectory tree generation

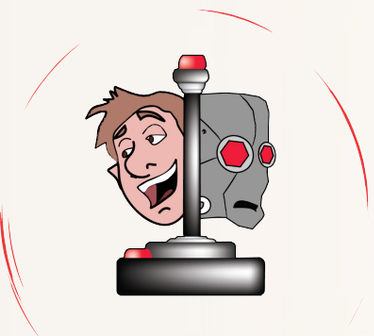
- Generate all possible plot sequences from a plot graph
- Including stories hypothesized to be possible based on the plot graph but not part of the exemplars



Semantic lowering

- Map plot events to agent operators
- Summed word2vec embeddings and cosine similarity





Agent actions:

0.83 → give_money

0.14 → pick_up_prescription

enter_pharmacy

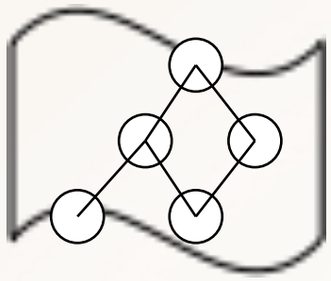
exit_pharmacy

0.67 → drop_money

Quixote

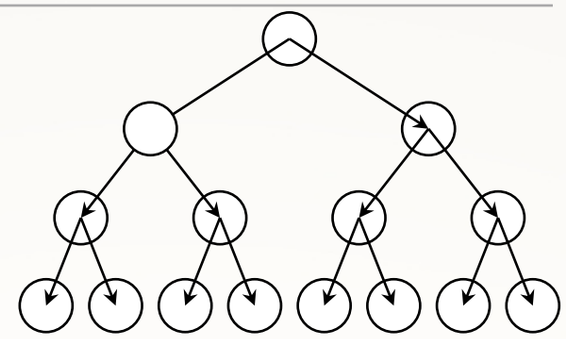
Exemplar stories

Plot graph learning



A plot graph

Trajectory tree creation

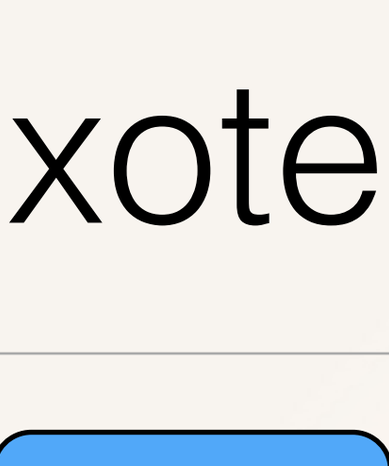


A trajectory tree

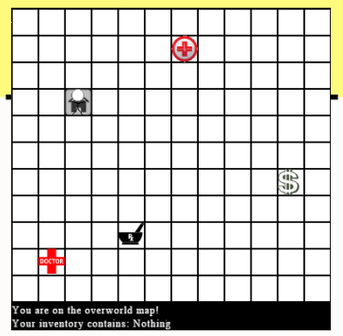
Reward assignment

A trajectory tree with events assigned reward values

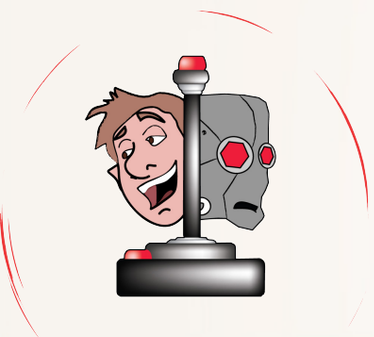
Reinforcement learning



A policy mapping states to actions

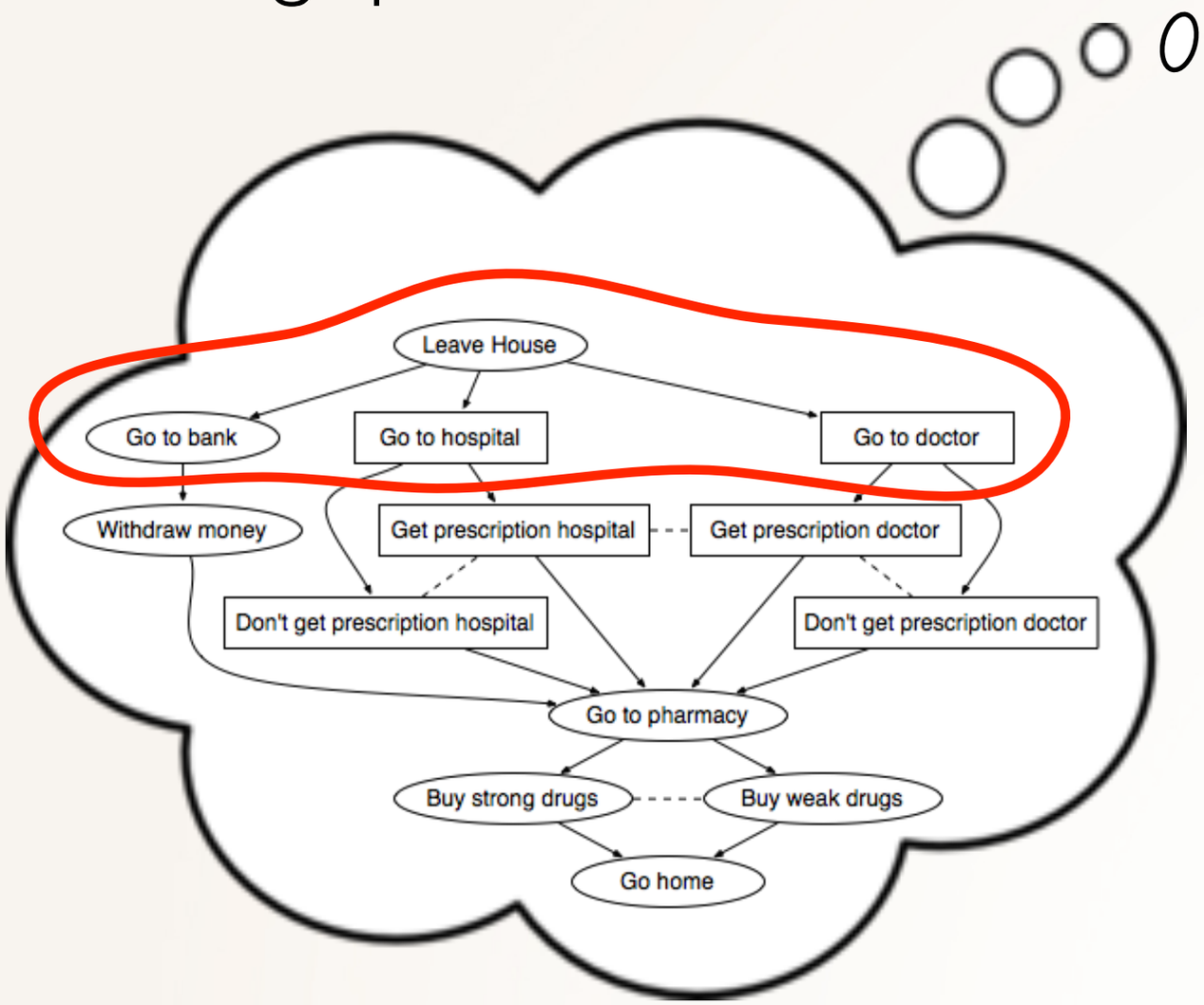


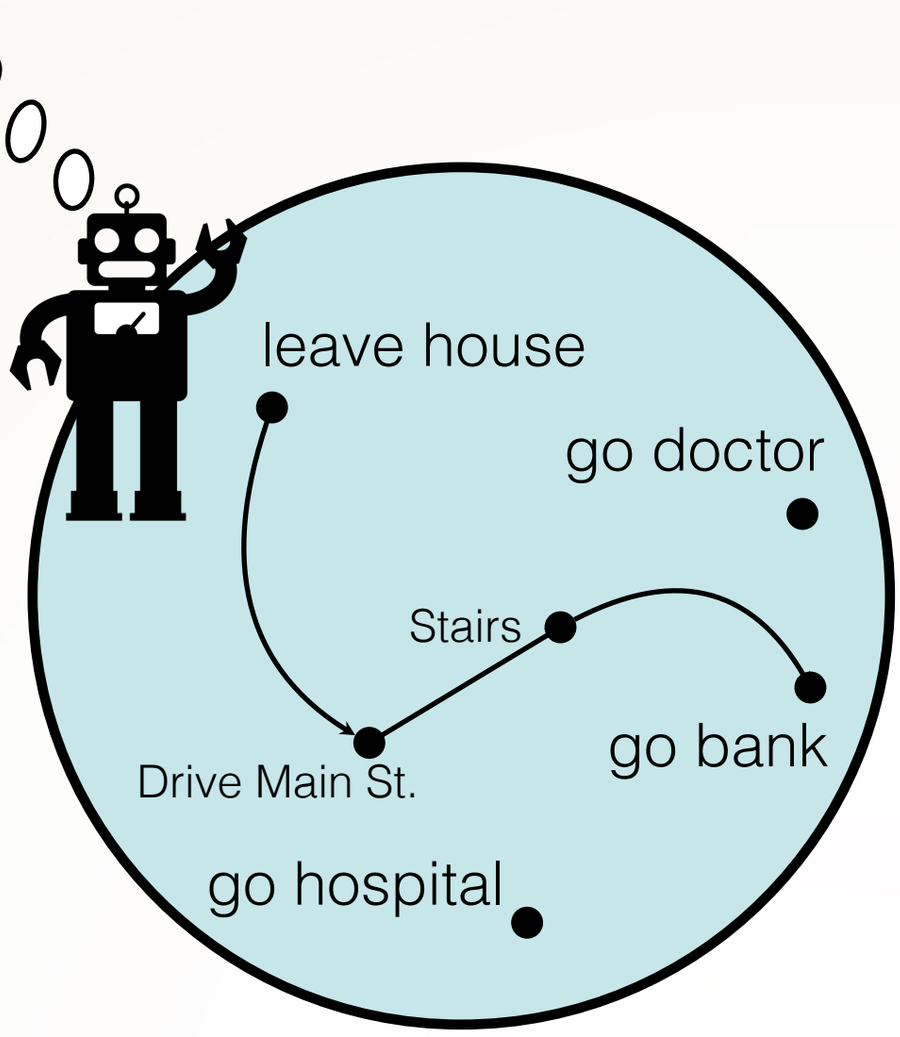
Environment



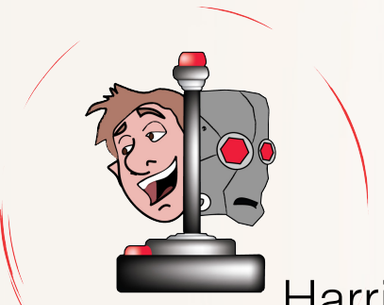
Reinforcement learning

- Fill gaps between events





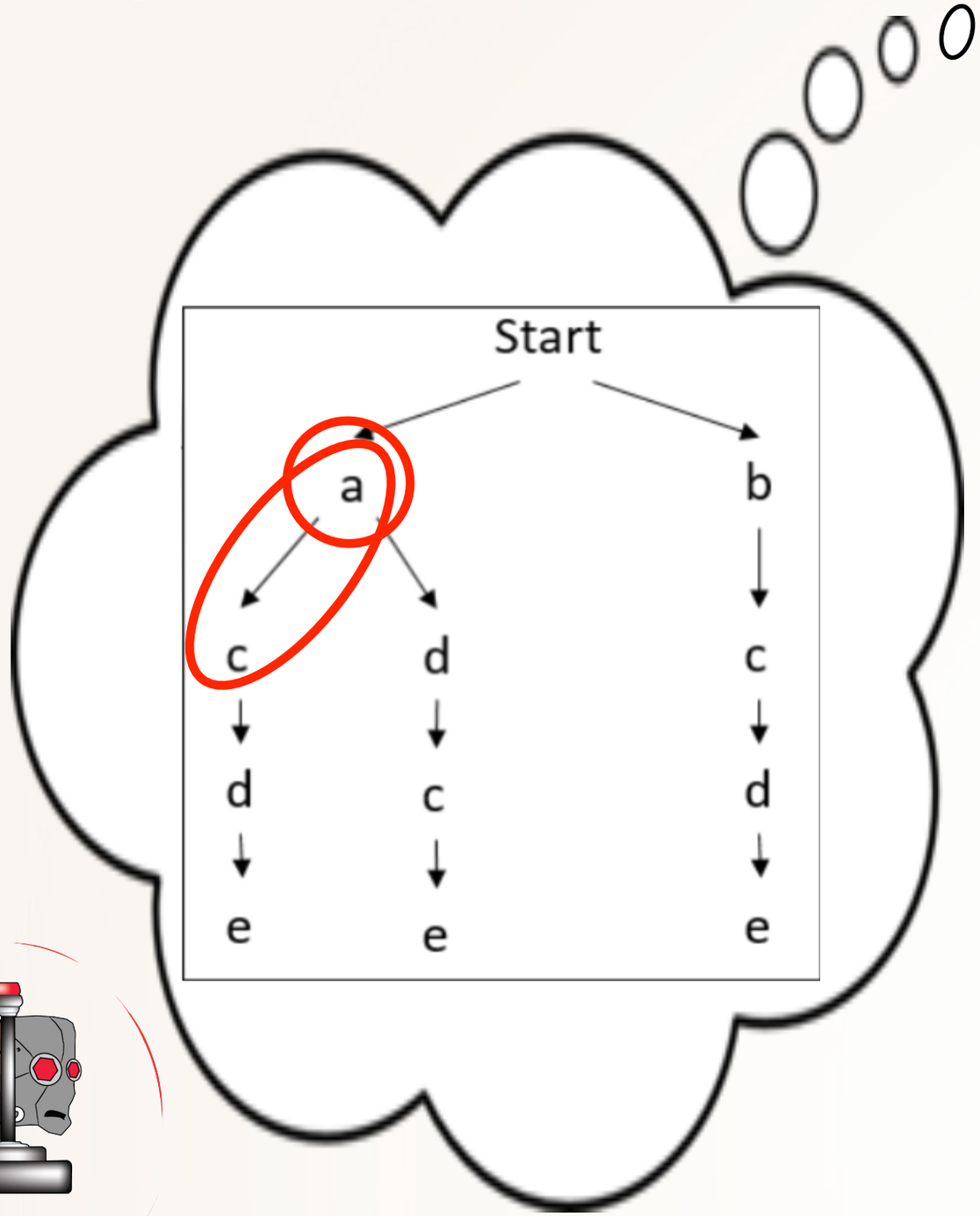
World state space

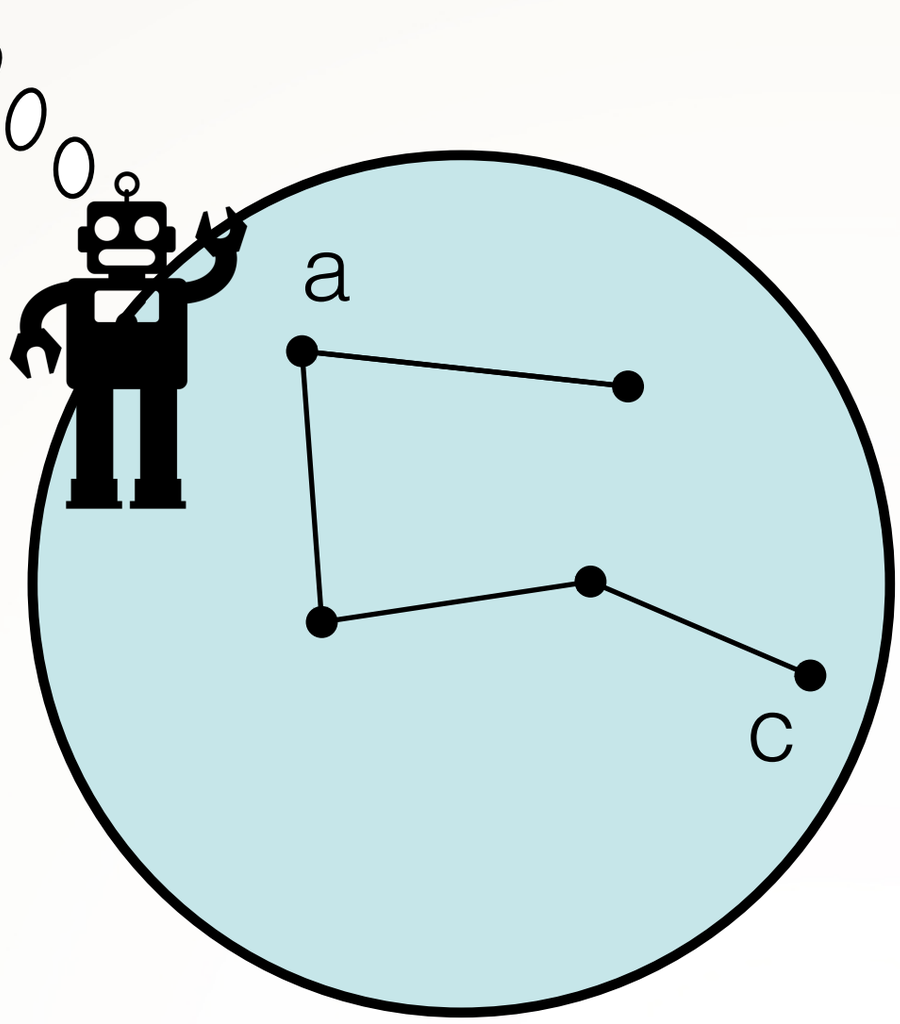


Harrison & Riedl. AIIDE Conference, 2016.

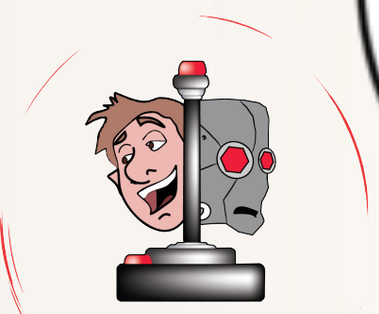
Reinforcement learning

- Fill gaps between events in trajectory tree





World state space



State and reward

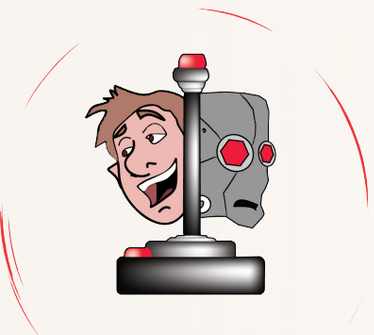
- Stories are non-Markovian but reinforcement learners make use of the Markov assumption

- Roll trajectory tree into state:

$$S' = S \times T$$

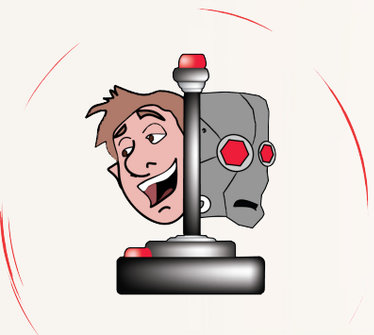
for all states in S
and trajectories in T

- Agent receives a reward every time it performs an action that advances the trajectory tree; otherwise penalty



Reinforcement learning

- Q-learning with ε-greedy strategy
- Reinforcement learning not beholden to trajectory tree
- Fills gaps between trajectory tree events



Pharmacy world

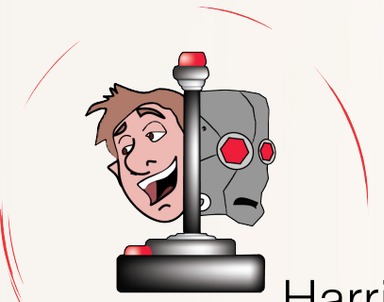
NEXT WE TEACH QUIXOTE SOCIAL CONVENTIONS BY TELLING IT STORIES ABOUT GOING TO THE PHARMACY

Human subjects study

- Which agent is most humanlike?

- Conditions:
  - 1: No failures
  - 2: One failure
  - 3: Two failures

	Quixote	Baseline	Both
Condition 1	23	2	7
Condition 2	28	10	4
Condition 3	23	8	3



Harrison & Riedl. IJCAI Workshop on Interactive ML, 2016.

AI safety

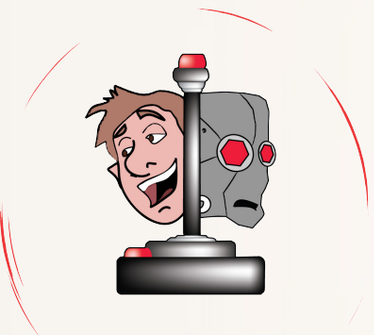
- Use stories to correct for commonsense goal errors by "filling in" incomplete reward functions

- Supporting theory of mind: agent acts more in line with user expectations

- Plot graphs learned from stories can be used by AI agents to predict behaviors by humans

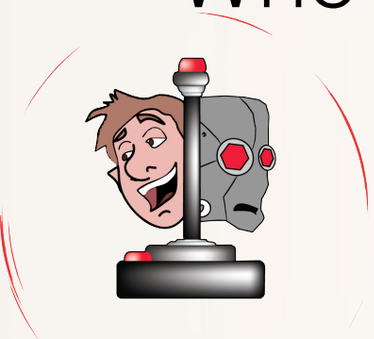
Purdy & Riedl. Reading Between the Lines. ICIDS Conference, 2016

Plot Graph	Graph Size (# Nodes)	Accuracy	# of Brute-Force String Matches	# of Heuristic String Matches
Restaurant	19	0.7395	430,931,917	54,060,547
Bank Robbery	28	0.6980	5,705,264,000	285,024,000



Next steps

- Quixote is a proof of concept
- Crowdsourced stories —> story corpora in the wild
- Deep reinforcement learning
  - Scale up plot graph learning
  - Correspondence between sentences and effectors
- Different agents for different cultures
- Who decides what the agent reads?



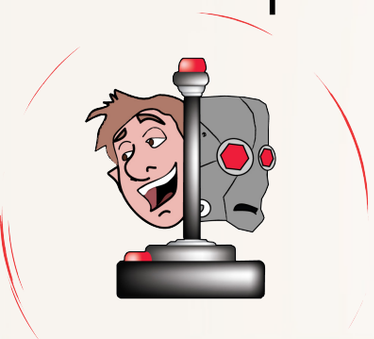
Concluding thoughts

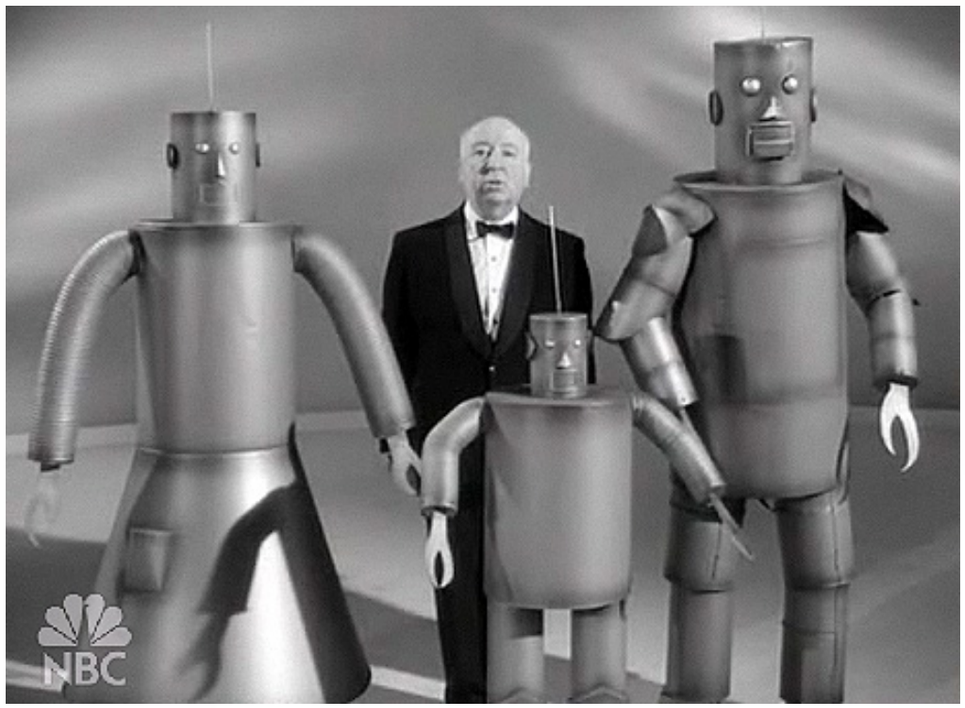
- Narrative intelligence is central to many of the things humans do on a day to day basis

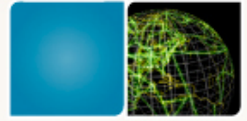
- Computational narrative intelligence is a key capability that unlocks many practical, real-world applications

- Computational narrative intelligence is about understanding humans and appearing less alien

- A path to “safer” AI

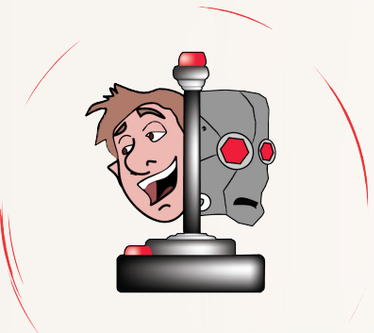




Thanks!

- Brent Harrison
- Boyang (Albert) Li
- Stephen Lee-Urban
- Siddhartha Banerjee
- D. Scott Appling
- George Johnston
- Yijie Wang



http://www.cc.gatech.edu/~riedl/

riedl@gatech.edu @mark_riedl

Georgia Tech College of Computing
School of Interactive Computing