Chatbots from First Principles
AVIOS Conference
Feb. 5-6, 2018
Tim Connelly/Jonathan Mugan, PhD

Siliconic Home
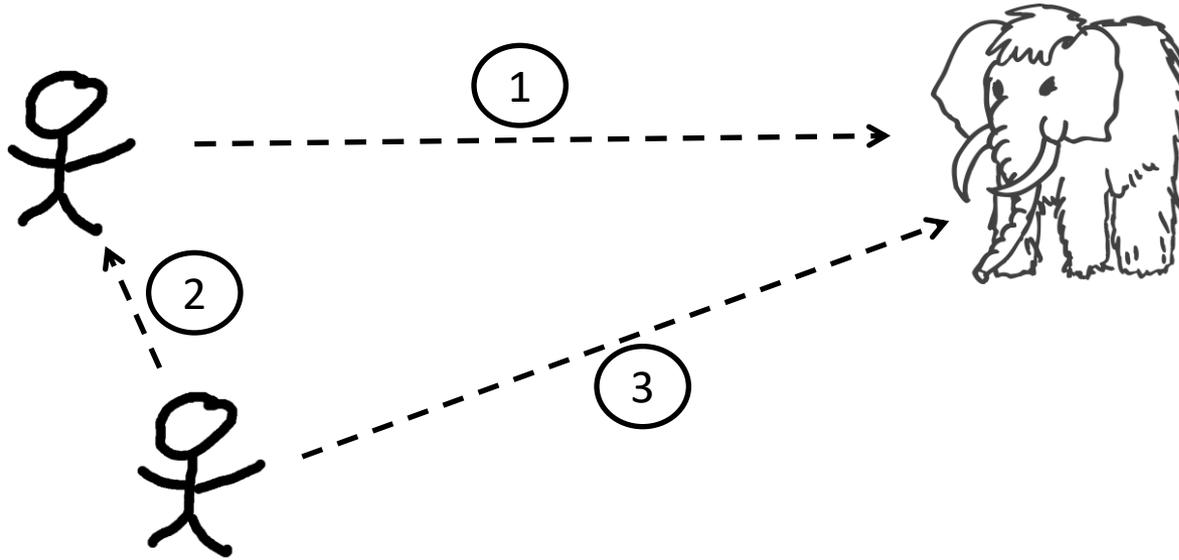
Talk outline

What conversation is

- The current state of chatbots
- Intention-based agents
- Conversational agents

Conclusion

Siliconic Home

Conversation requires shared reference

Language begins with shared attention by pointing to things in the world. Words then point to shared ideas in our minds [Gärdenfors, 2014].
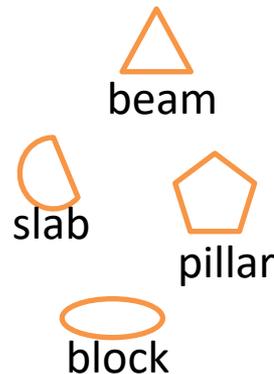
Language as shared convention

Wittgenstein and his language games,
Philosophical Investigations, 1958
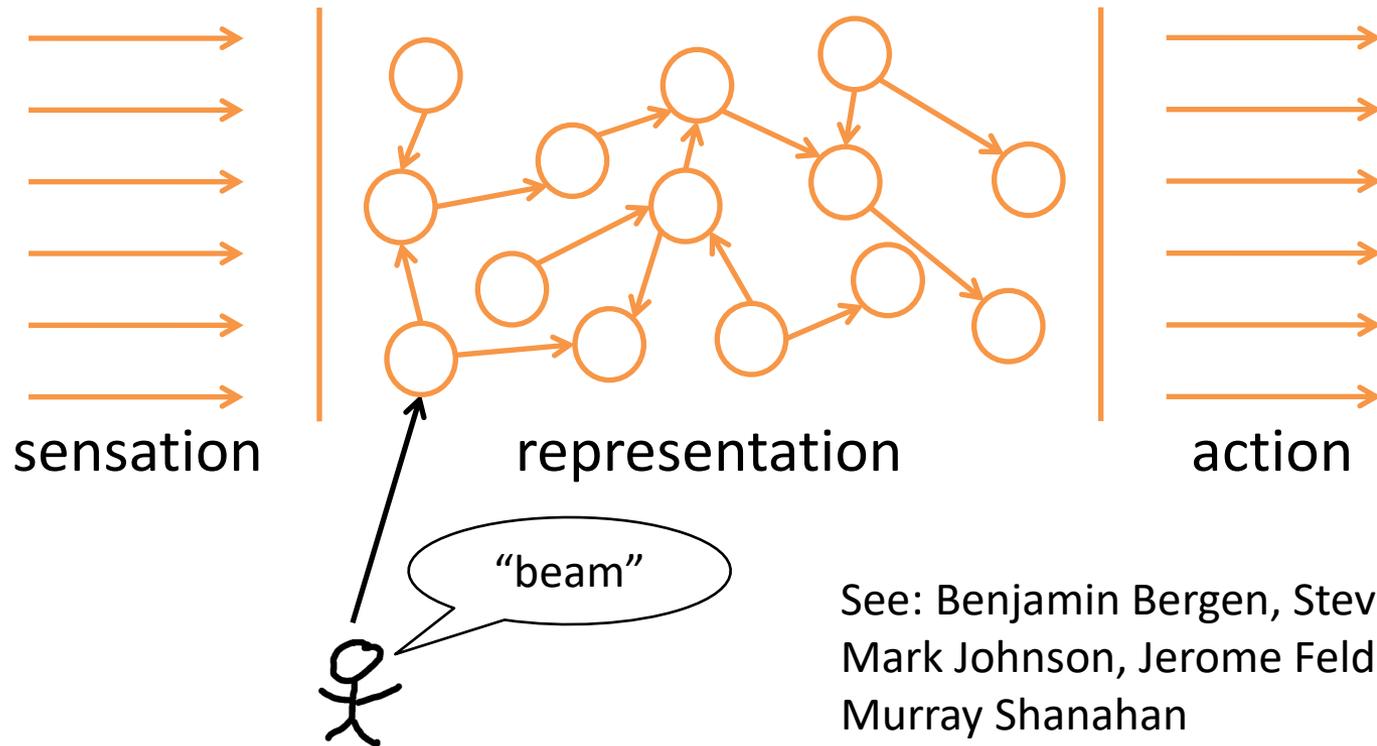
Two people building something.
"Bring me a beam."

A

beam

slab

pillar

block

B

Eventually turns into a shared convention for a community

Siliconic Home

Our brains map community conventions to personal sensations and actions

This is what is means for language to be grounded.

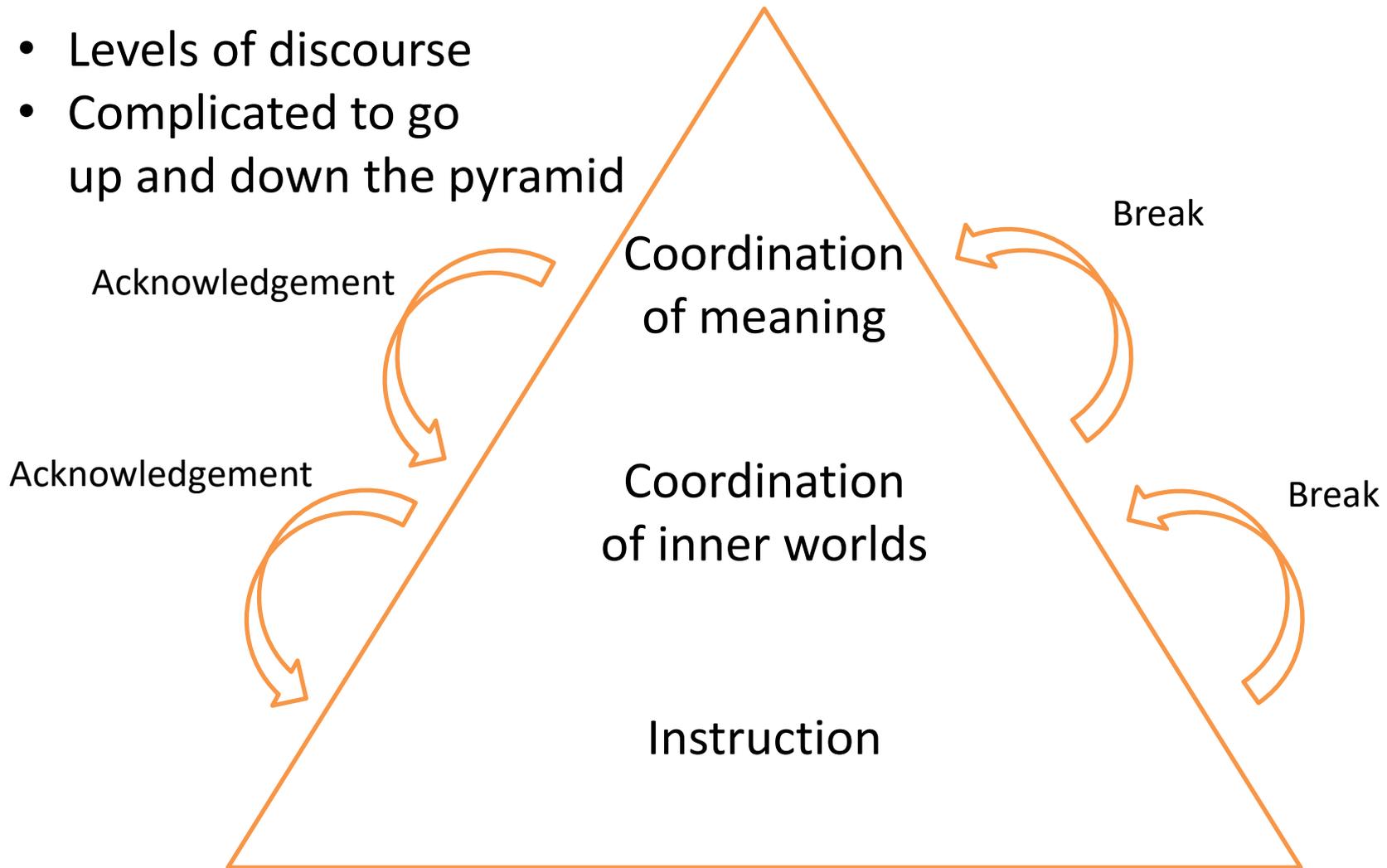sensation   representation   action

"beam"

See: Benjamin Bergen, Steven Pinker, Mark Johnson, Jerome Feldman, and Murray Shanahan

- When someone says "beam," we map that to our experience with beams.
- We understand each other because we have had similar experiences.

We negotiate language and meaning as we go

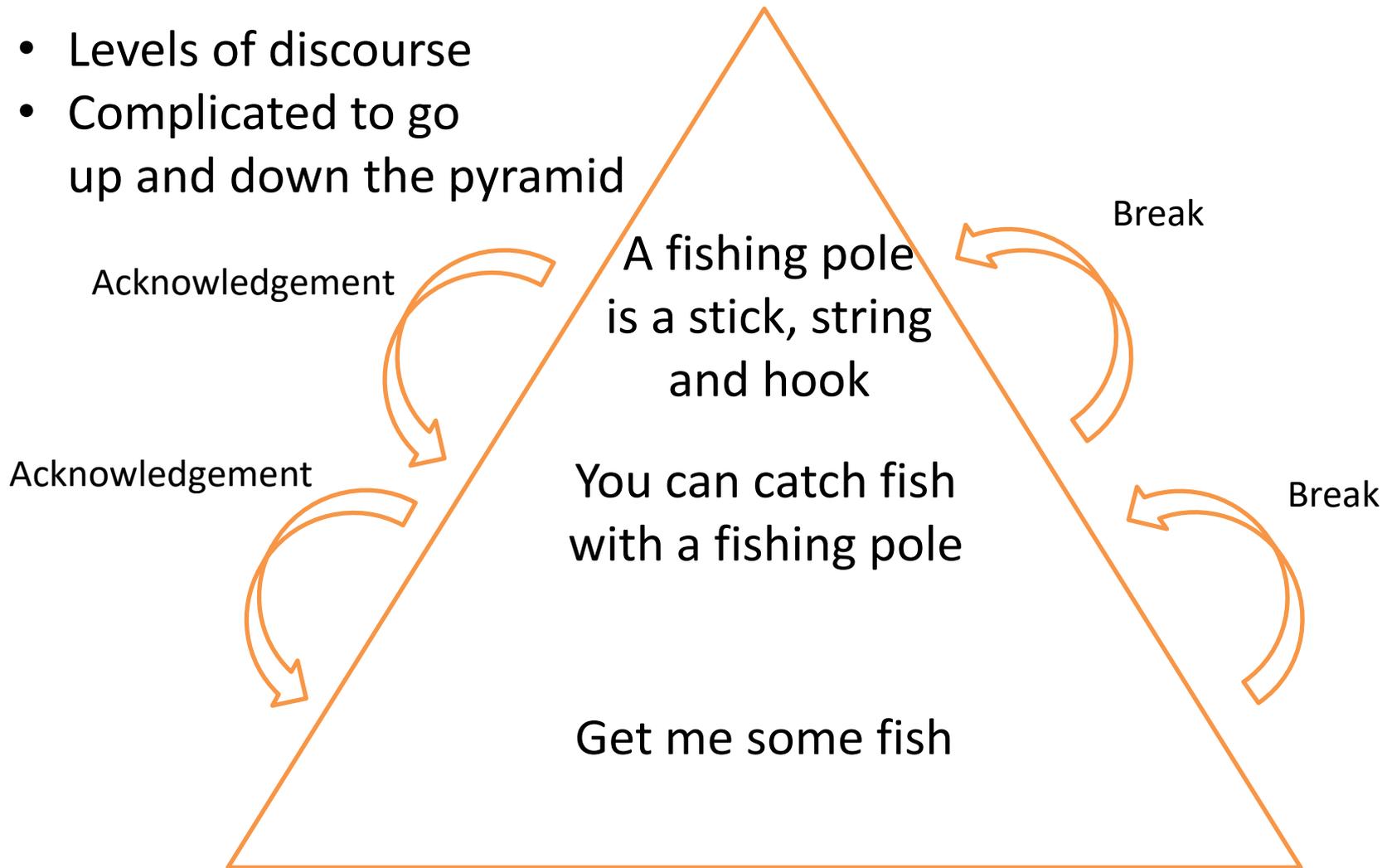- Levels of discourse
- Complicated to go
  up and down the pyramid

Acknowledgement

Acknowledgement

Break

Break

Coordination
of meaning

Coordination
of inner worlds

Instruction

Modified from Gärdenfors (2014), which was based on Winter (1998)

We negotiate language and meaning as we go

- Levels of discourse
- Complicated to go
  up and down the pyramid

Acknowledgement

Acknowledgement

Break

Break

A fishing pole
is a stick, string
and hook

You can catch fish
with a fishing pole

Get me some fish

Modified from Gärdenfors (2014), which was based on Winter (1998)

Conversation has its own rules (pragmatics)

- Conversational maxims: Grice (1975, 1978)
- Breaking these rules is a way to communicate more than the meaning of the words.

Maxim of Quantity: Say only what is not implied.
Yes: Bring me the block.
No: Bring me the block by transporting it to my location.
What did she mean by that?

Maxim of Quality: Say only things that are true.
Yes: "I hate carrying blocks."
No: "I love carrying blocks, especially when they are covered in fire ants."
She must be being sarcastic.

Maxim of Relevance: Say only things that matter.
Yes: "Bring me the block."
No: "Bring me the block and birds sing."
What did she mean by that?

Maxim of Manner: Speak in a way that can be understood.
Yes: "Bring me the block."
No: "Use personal physical force to levitate the block and transport it to me."
What did she mean by that?

We've all heard of Eliza

Simple substitutions to mimic a psychologist from the 1960s

"My mother wants me to buy a bazooka." …
"Tell me why your mother wants you to buy a bazooka."

You can extend it as much as you want

I want to buy a car. → How much does a car cost?

Siliconic Home

Intention-based agents

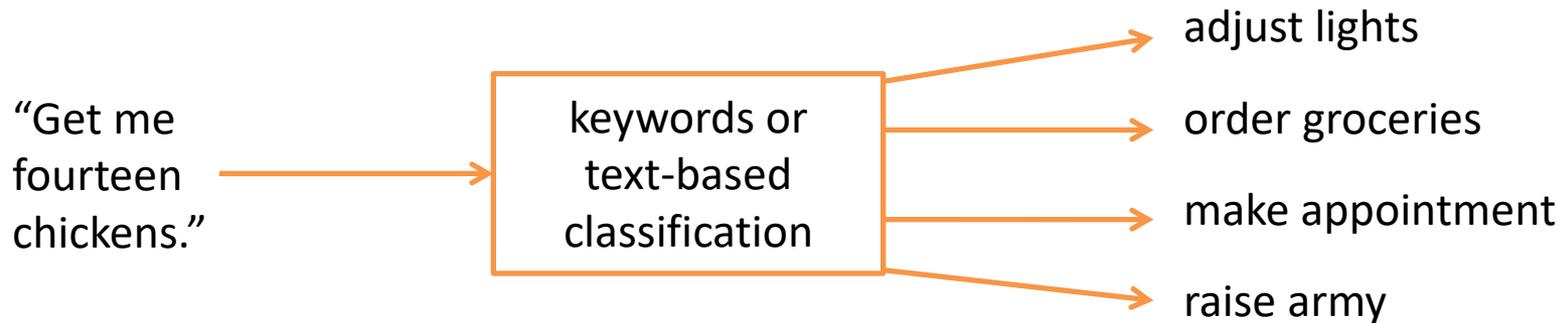Examples: Amazon Echo, Google Assistant, Siri, Cortana, Apple

Like a command language
1. Identify what the user wants the machine to do (the "intent")
2. Figure out the details of the intent so the machine can take action

Siliconic Home

How to determine the intent?

Use keywords or text-based classification

"Get me fourteen chickens." → keywords or text-based classification → adjust lights

order groceries

make appointment

raise army

With a bag-of-words in scikit-learn (library for Python)
http://scikit-learn.org/stable/auto_examples/applications/plot_out_of_core_classification.html

Using deep learning with a convolutional neural network (CNN) in TensorFlow (use this if you have a lot of data)
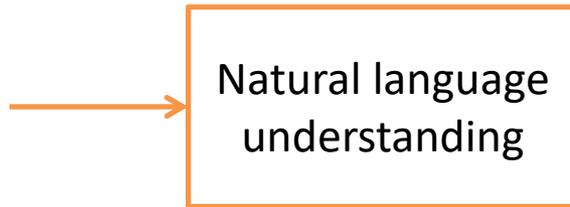https://github.com/dennybritz/cnn-text-classification-tf

Siliconic Home

What to do with the intent

Convert the squishiness of language into a Python dictionary

"Get me fourteen chickens." → Natural language understanding → {'domain': 'purchase', 'item':'chicken_id344', 'quantity': 14}

Called a frame and slot semantics

Siliconic Home

Natural language understanding with context free grammars (compositional semantics)

s[0] is the semantic value of the first item on rule right hand side
s[1] is the semantic value of the second item on rule right hand side


```
$Order → $Purchase $ItemAmount, dunion(s[0],s[1])
$Purchase → is_purchase(tokens), {'domain': 'purchase'}
$ItemAmount → $Amount $Item, dunion(s[0],s[1])
$Amount → is_number(tokens), {'amount': get_number(tokens)}
$Item → is_item(tokens), {'item': get_item(tokens)}
```

```
def is_purchase(tokens):
    return tokens in ['get me', 'buy', 'grab me some']
```

```
get_item('chickens') → 'chicken_id344'
get_item('soap') → 'soap_id143'
get_number(fourteen) → 14
```

```
dunion: dictionary union
```

Siliconic Home

Conversational agents

Extended, meaningful conversations. Have to be able to keep track of the state of the conversation and know when the person wants to talk about something else.

Blog post I wrote, You and Your Bot: A New Kind of Lifelong Relationship
https://chatbotsmagazine.com/you-and-your-bot-a-new-kind-of-lifelong-relationship-6a9649feeb71

A teacher for the young
- Starts out as a cell phone app for a child
- Child can talk to its cartoon face and show it her toys
- Teaches her things
    - "If you have 7 giraffes and bought 3 more, how many would you have?"
    - "If you eat your chocolate bear today, how are you going to feel tomorrow when you no longer have it?"

A guide for the adult
- As the child grows, so does the app, so it becomes her operating system
- Knows what you know. Can give turn-by-turn directions on fixing the washing machine

An advocate for the old
- Could advocate for us and guide us through government and healthcare bureaucracies
- Could help us live independently longer.
    - Imagine it sees you confused. Could talk you through making coffee.

How could a machine manage such conversations?

A dialog manager needs to

Take the person through all the things it wants to talk about
- Teach her math
- Then talk about planets

Recognize when she wants to talk about something else
and queue the current topics for later

RavenClaw from CMU is probably the best known dialog manager

Siliconic Home

Hierarchy of dialog agents

Ask 4+5
Teach Arithmetic
Teach Planets
Tell Stories

Dialog Stack

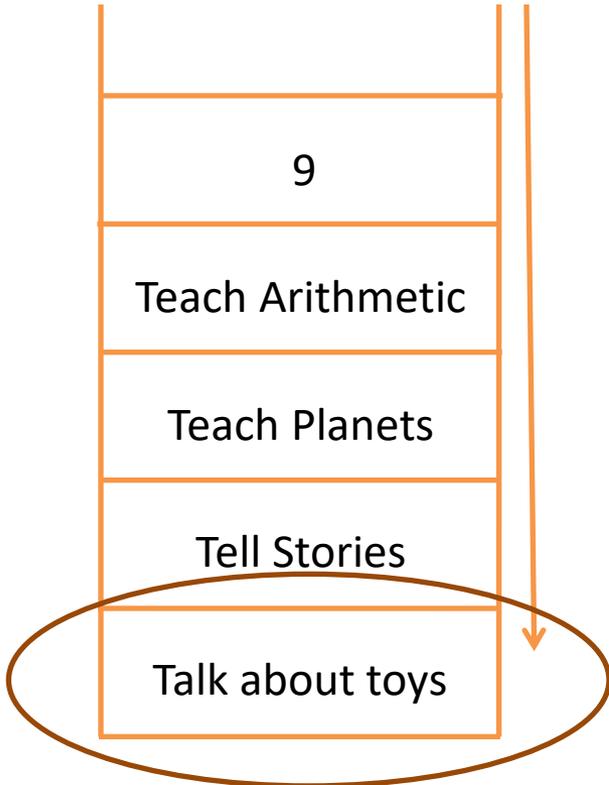| 9 |
|---|
| Teach Arithmetic |
| Teach Planets |
| Tell Stories |
| Talk about toys |

Expectation Agenda

Talk about toys always there.

Siliconic Home

Hierarchy of dialog agents

Robot just finished asking, "What is 4 + 5?"
Child says: "Do you like Mr. Fluffles?"

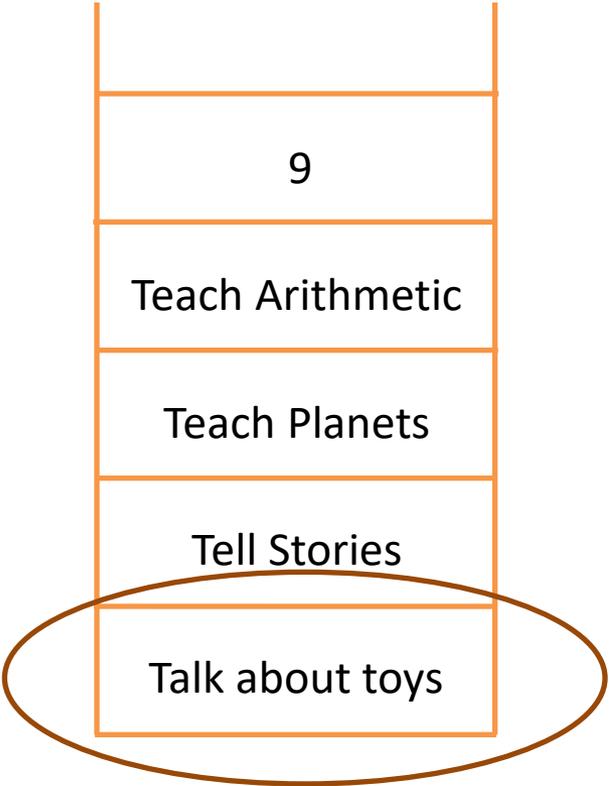| Dialog Stack	Expectation Agenda
Ask 4+5	9
Teach Arithmetic	Teach Arithmetic
Teach Planets	Teach Planets
Tell Stories	Tell Stories
	Talk about toys

Siliconic Home

Hierarchy of dialog agents

Robot just finished asking, "What is 4 + 5?"
Child says: "Do you like Mr. Fluffles?"
Robot responds: "Very nice. Is he your favorite?"

Dialog Stack
Ask 4+5
Teach Arithmetic
Teach Planets
Tell Stories

Expectation Agenda
9
Teach Arithmetic
Teach Planets
Tell Stories
Talk about toys

Siliconic Home

Hierarchy of dialog agents

Robot just finished asking, "What is 4 + 5?"
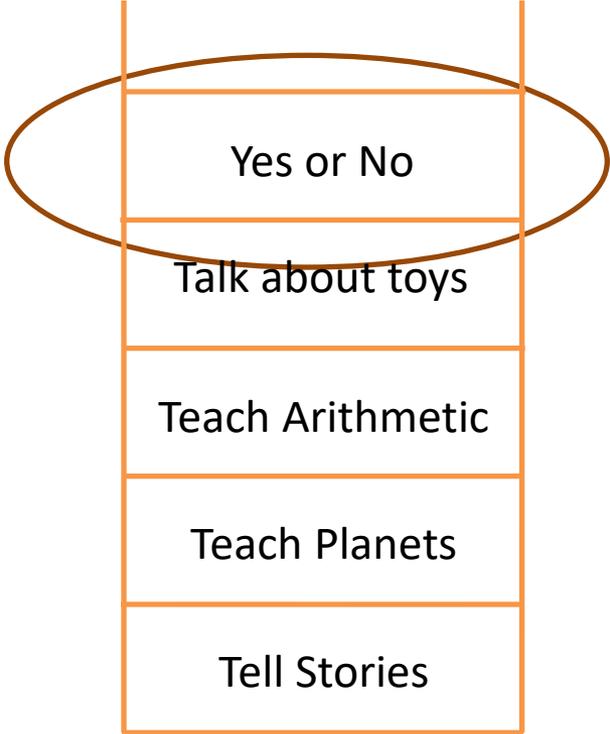Child says: "Do you like Mr. Fluffles?"
Robot responds: "Very nice. Is he your favorite?"

Now we are talking about toys.

Dialog Stack
Talk about toys
Teach Arithmetic
Teach Planets
Tell Stories

Expectation Agenda
Yes or No
Talk about toys
Teach Arithmetic
Teach Planets
Tell Stories

Siliconic Home

Hierarchy of dialog agents

Robot just finished asking, "What is 4 + 5?"
Child says: "Do you like Mr. Fluffles?"
Robot responds: "Very nice. Is he your favorite?"

Child says, "Of course!"
Update database
with new favorite toy.

Now we are talking about toys.

| Talk about toys |
| Teach Arithmetic |
| Teach Planets |
| Tell Stories |

Dialog Stack

| Yes or No |
| Talk about toys |
| Teach Arithmetic |
| Teach Planets |
| Tell Stories |

Expectation Agenda

Siliconic Home

State-based dialog agents

Markov decision process
Set of **states**
Set of **actions**
Get a **reward** for being in a state s and taking an action a

States are things such as what the bot knows (questions it has answered),
the last thing the bot said, and the last thing the user said.

Actions are making particular statements.

Reward comes from meeting a goal state, such as child giving correct answer to a math problem or successfully completing a travel reservation.
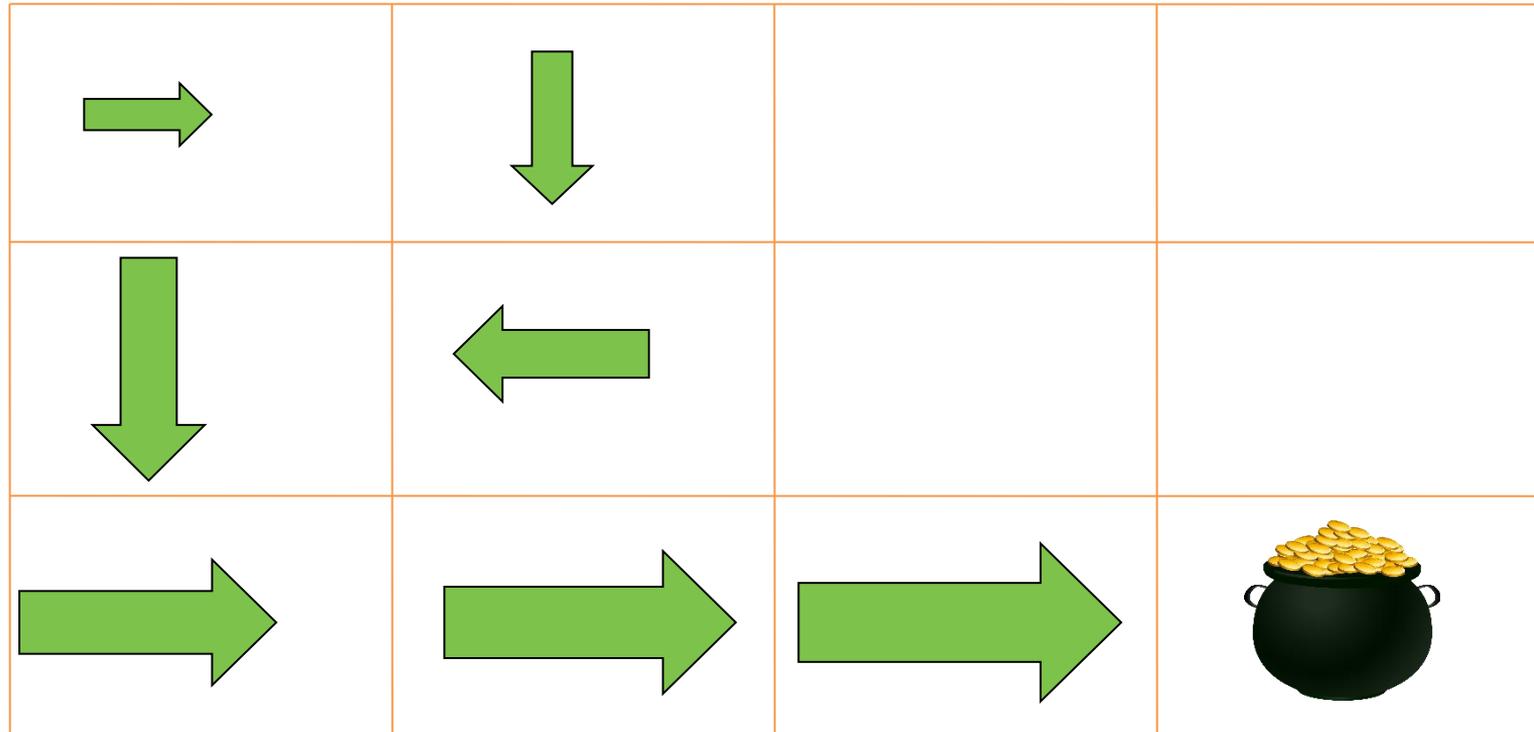
Use **reinforcement learning** to learn a policy that gives the best action a for being in state s.

Siliconic Home

# Beginning with random exploration
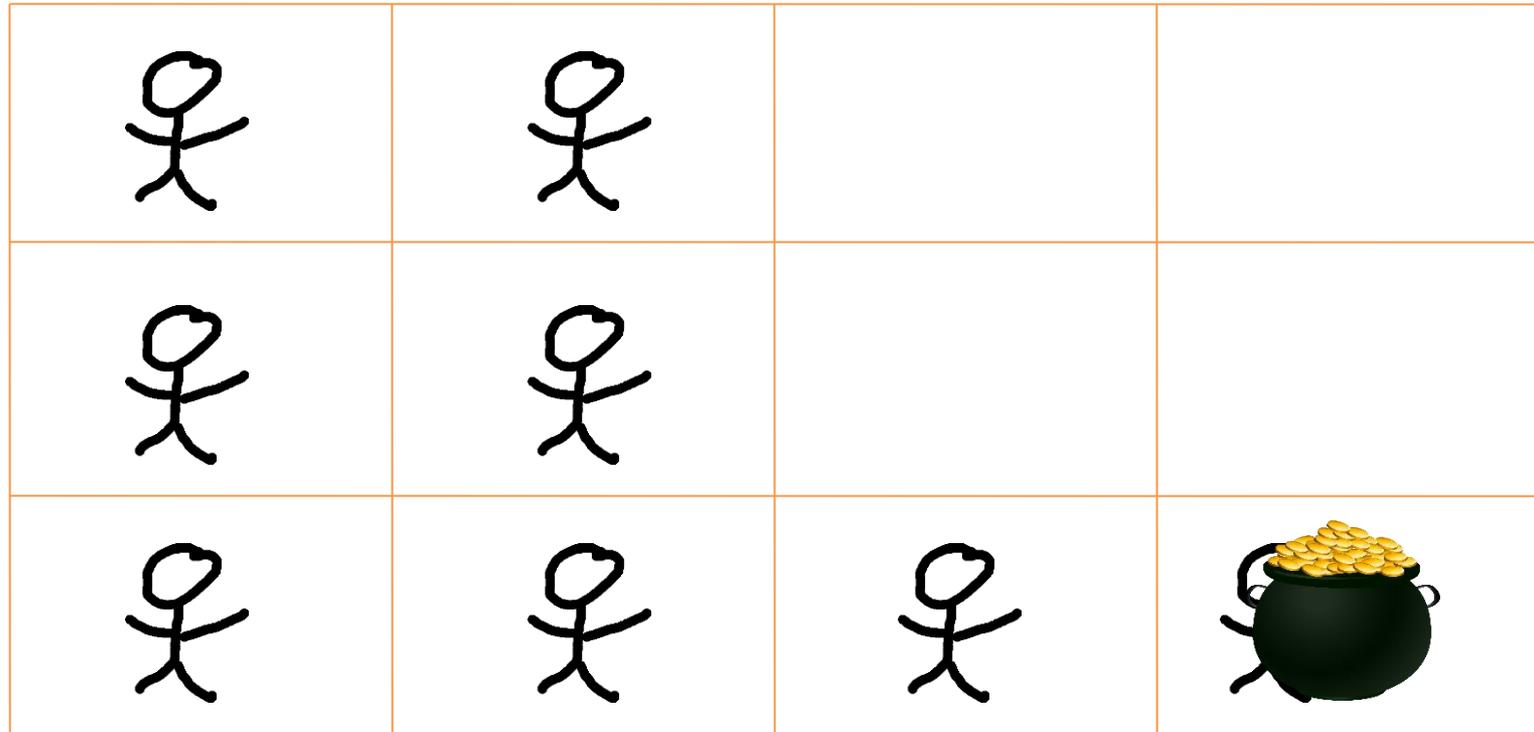
In reinforcement learning, the agent begins by randomly exploring until it reaches its goal.

Siliconic Home

Reaching the goal

- When it reaches the goal, credit is propagated back to its previous states.
- The agent learns the function $Q(s,a)$, which gives the cumulative expected discounted reward of being in state s and taking action a and acting according to the policy thereafter.

Learning the behavior
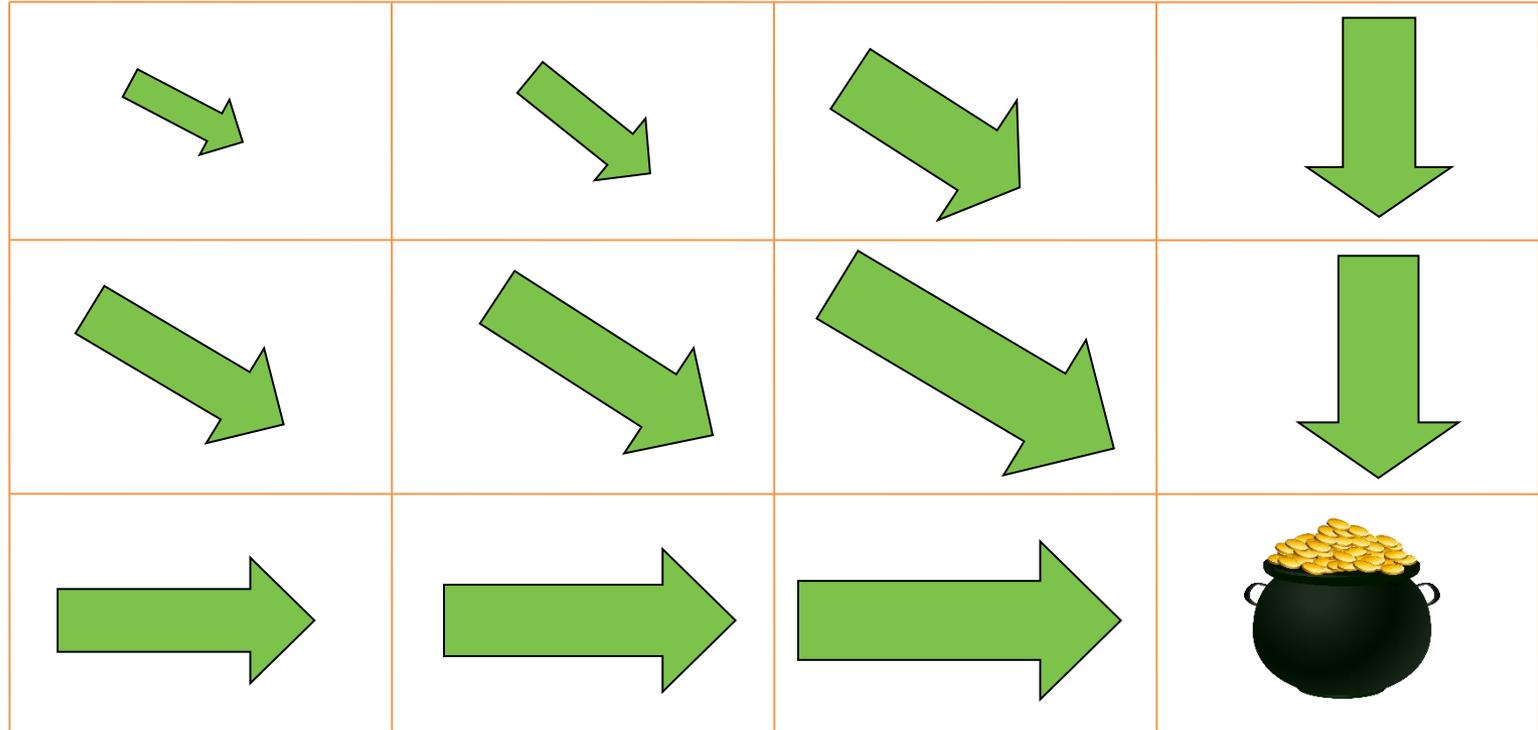
Eventually, the agent learns the value of being in each state and taking each action and can therefore always do the best thing in each state.

# But there is limited observability

- Since speech-to-text and natural language understanding are error prone, you actually don't know for sure which state you are in.

- Does she want to talk about her toys, or is she telling me that I should be fluffier?

- Partially Observable Markov Decision Process (POMDP)

- As you can guess, these need a lot of training data. Generally have to create a simulation of a person for the bot to talk to.

Siliconic Home

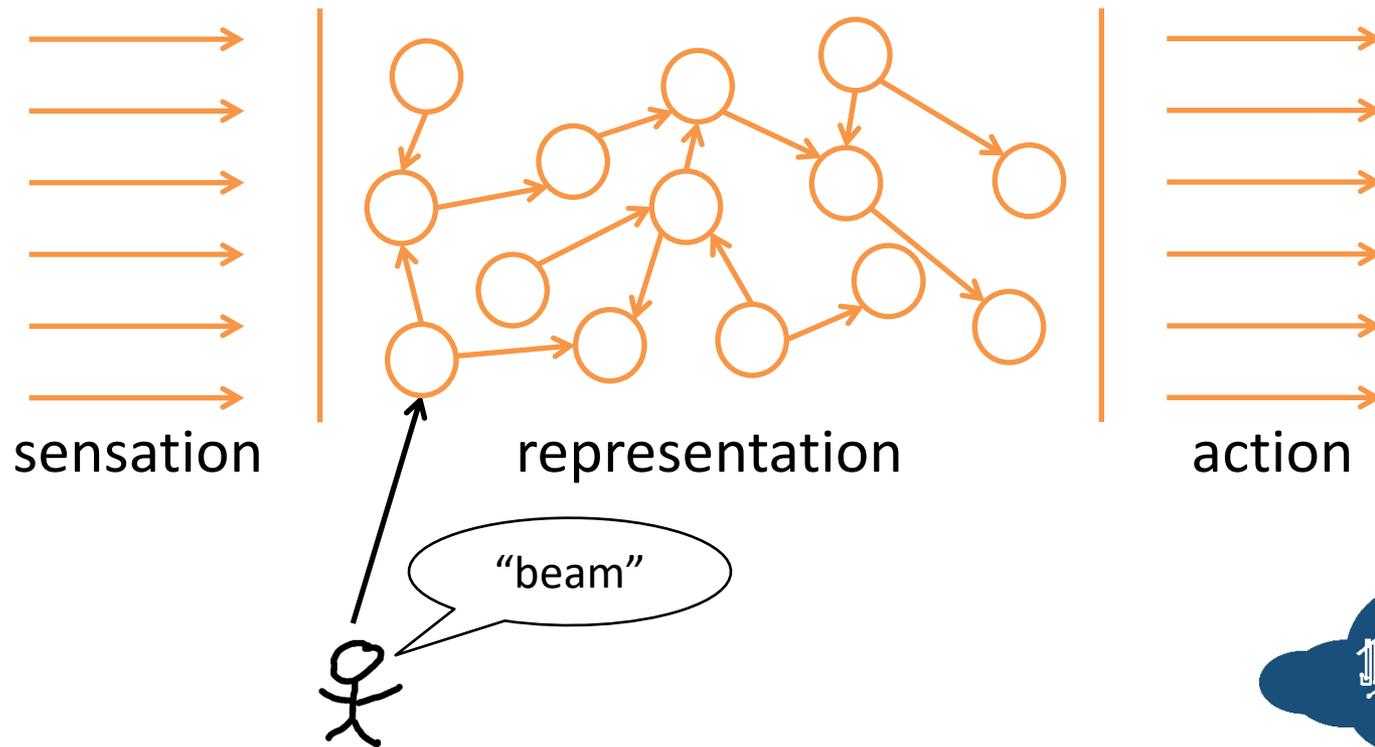Computers need to know our conventions

A

beam

slab

pillar

block
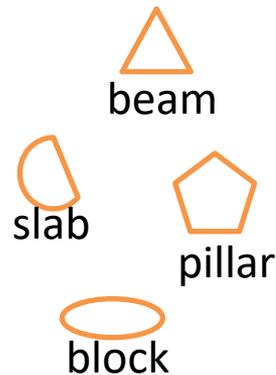
B

We can program this in

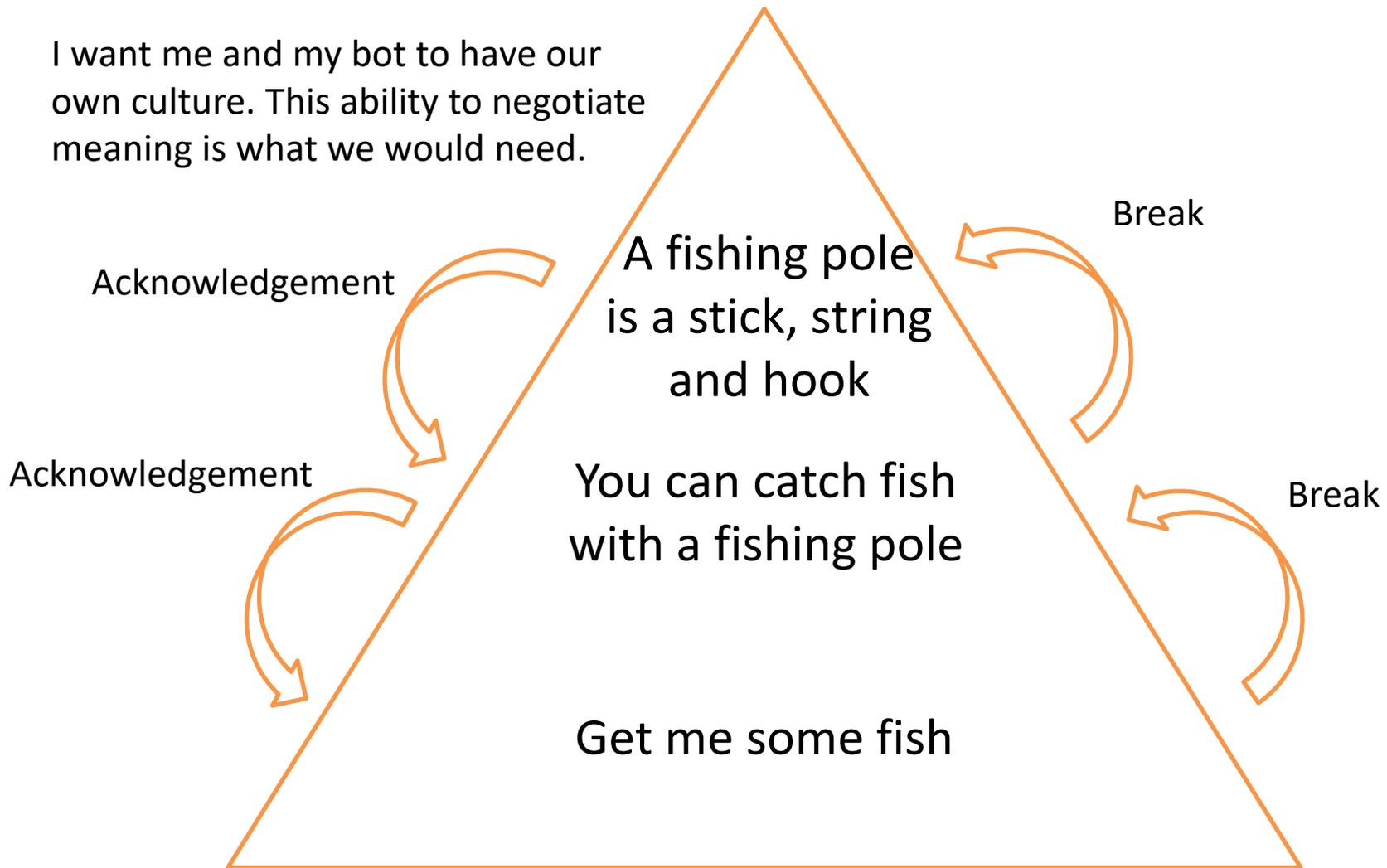Siliconic Home

- Conventions need to be grounded in a system of sensation and action
- Hard when machines don't have our bodies and experience
- Blog post of wrote in using physics simulations for natural language understanding https://chatbotslife.com/computers-could-understand-natural-language-using-simulated-physics-26e9706013da

I want me and my bot to have our own culture. This ability to negotiate meaning is what we would need.

Break

Acknowledgement

A fishing pole is a stick, string and hook

Acknowledgement

You can catch fish with a fishing pole

Break

Get me some fish

Machines also need to know the rules of conversation, so they can understand meanings beyond words. And they need to understand meanings from how words are said.

Thanks for listening

Questions?

tconnelly@siliconichome.com

Siliconic Home