



SI 630

Natural Language Processing: Algorithms and People

Lecture 12: Dialog
March 18, 2018

All the students on their webcams
waiting for class to start



Some slides adapted from David Bamman

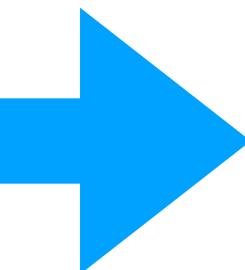
Administration time!

Class Schedule B.C.

Week #	Assigned	Due
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		Project Update
12	HW5	HW4
13		Midterm
14		Project Presentation
15		HW5
Finals Week		Project Report

Class Schedule B.C.

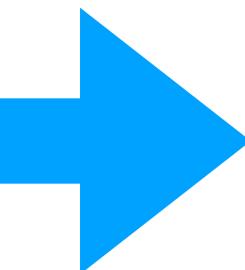
Week #	Assigned	Due
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		Project Update
12	HW5	HW4
13		Midterm
14		Project Presentation
15		HW5
Finals Week		Project Report



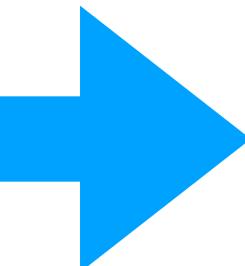
Class Schedule B.C.

Class Schedule A.D. (new one)

Week #	Assigned	Due
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		Project Update
12	HW5	HW4
13		Midterm
14		Project Presentation
15		HW5
Finals Week		Project Report



Class Schedule B.C.

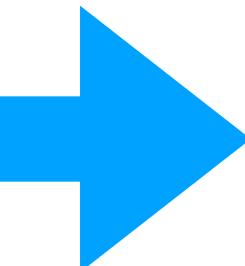


Week #	Assigned	Due
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		Project Update
12	HW5	HW4
13		Midterm
14		Project Presentation
15		HW5
Finals Week		Project Report

Class Schedule A.D. (new one)

Week #	Assigned	Nominal Due Date
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		
12	HW5	HW4 + Project Update
13		Midterm
14		
15		HW5
Finals Week		Project Report + Project Blogpost

Class Schedule B.C.



Week #	Assigned	Due
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		Project Update
12	HW5	HW4
13		Midterm
14		Project Presentation
15		HW5
Finals Week		Project Report

Class Schedule A.D. (new one)

Week #	Assigned	Nominal Due Date
9	Relaxing (Spring Break)	Relaxing
10	HW4	HW3
11		
12	HW5	HW4 + Project Update
13		Midterm
14		
15		HW5
Finals Week		Project Report + Project Blogpost

Regarding due dates

Regarding due dates

- All homework and project due dates are now nominal

Regarding due dates

- All homework and project due dates are now nominal
- You can turn HW4 and HW5 in at any time until the end of the semester

Regarding due dates

- All homework and project due dates are now nominal
- You can turn HW4 and HW5 in at any time until the end of the semester
- **One exception:** Project Update must be turned in by April 15

Regarding due dates

- All homework and project due dates are now nominal
- You can turn HW4 and HW5 in at any time until the end of the semester
- **One exception:** Project Update must be turned in by April 15
- Everything is still due though—don't let delayed work turn into delayed anxiety (but please reach out if you have concerns!)

Regarding due dates

- All homework and project due dates are now nominal
- You can turn HW4 and HW5 in at any time until the end of the semester
- **One exception:** Project Update must be turned in by April 15
- Everything is still due though—don't let delayed work turn into delayed anxiety (but please reach out if you have concerns!)
- Hard cutoff April 29 at 5:30pm—**We have to turn in grades within 24 hours of this**

What's in that Project Update?

What's in that Project Update?

- **Critical part**: A baseline system that is run on your data and evaluated using your metric

What's in that Project Update?

- **Critical part**: A baseline system that is run on your data and evaluated using your metric
- A plan for *your* method to solve your problem

What's in that Project Update?

- **Critical part**: A baseline system that is run on your data and evaluated using your metric
- A plan for *your* method to solve your problem
- Discussion of how well the baseline does

What's in that Project Update?

- **Critical part**: A baseline system that is run on your data and evaluated using your metric
- A plan for *your* method to solve your problem
- Discussion of how well the baseline does
- An updated **work plan**! Make a plan and stick to it

What's in that Project Update?

- **Critical part**: A baseline system that is run on your data and evaluated using your metric
- A plan for *your* method to solve your problem
- Discussion of how well the baseline does
- An updated **work plan**! Make a plan and stick to it

Breakout Opportunity #1: In small groups, describe your project, discuss your proposed baseline and ask for feedback (5 minutes)



words



morphology

words

syntax

morphology

words

discourse

syntax

morphology

words

semantics

discourse

syntax

morphology

words

pragmatics

semantics

discourse

syntax

morphology

words

This Week's Learning Goals

This Week's Learning Goals

- Think about text beyond the sentence boundary

This Week's Learning Goals

- Think about text beyond the sentence boundary
- Learn about how some words refer to earlier words, what to do about them, and how to computationally recognize this relationship

This Week's Learning Goals

- Think about text beyond the sentence boundary
- Learn about how some words refer to earlier words, what to do about them, and how to computationally recognize this relationship
- Understand what's the relationship between these sentences "Times were hard. They persevered." and how to computationally model this

This Week's Learning Goals

- Think about text beyond the sentence boundary
- Learn about how some words refer to earlier words, what to do about them, and how to computationally recognize this relationship
- Understand what's the relationship between these sentences “Times were hard. They persevered.” and how to computationally model this
- How to make chatbots stop saying “I don’t know what you’re talking about”



Co-reference resolution

Discourse

- Discourse covers linguistic expression **beyond the boundary of the sentence.**
 - Dialogues: the structure of turns in conversation
 - Monologues: the structure of entire passages, documents



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



LUKE

I'll never join **you!**

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to **your father.**

LUKE

He told me enough! It was **you** who killed **him.**

VADER

No. **I am your father**

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know
it to be true.

LUKE

No! No! No!



LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

Coreference resolution

- “Trump met Putin today; **he**’s the leader of the US.

Coreference resolution

Barack Hussein Obama II (聆听 */bə'ra:k hu'sein ou'bæ:mə/*; born August 4, 1961) is the 44th and current President of the United States, and the first African American to hold the office. Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School, where he served as president of the *Harvard Law Review*. He was a community organizer in Chicago before earning his law degree. He worked as a civil rights attorney and taught constitutional law at the University of Chicago Law School from 1992 to 2004. He served three terms representing the 13th District in the Illinois Senate from 1997 to 2004, running unsuccessfully for the United States House of Representatives in 2000.

Coreference resolution

attend graduate school at [Harvard University](#) on a scholarship. Obama's parents divorced in March 1964.^[11] Obama Sr. returned to Kenya in 1964 where he remarried; he visited Barack in Hawaii only once, in 1971.^[12] He died in an automobile accident in 1982 when his son was 21 years old.^[13]

Did Barack Obama die in an automobile accident in 1982?

Coreference resolution

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

The entities or individuals **in the real world** that text is pointing to.

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

The entities or individuals **in the real world** that text is pointing to.

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

The entities or individuals **in the real world** that text is pointing to.

- VICTORIA CHEN

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

The entities or individuals **in the real world** that text is pointing to.

- VICTORIA CHEN
- MEGABUCKS

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referent”

The entities or individuals **in the real world** that text is pointing to.

- VICTORIA CHEN
- MEGABUCKS
- LOTSABUCKS

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referring expression”

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referring expression”

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Referring expression”

The text that points to entities.

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw her pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“coreference”

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw [her](#) pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“coreference”

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw [her](#) pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“coreference”

The set of text strings
that all refer to the same
ENTITY.

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw [her](#) pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“coreference”

The set of text strings that all refer to the same ENTITY.

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw [her](#) pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Coreference

“coreference”

The set of text strings that all refer to the same ENTITY.

“Victoria Chen, Chief Financial Officer of Megabucks Banking Corp since 2004, saw [her](#) pay jump 20%, to \$1.3 million, as the 37-year-old also became the Denver-based financial services company’s president. It has been ten years since she came to Megabucks from rival Lotsabucks.”

Event coreference

I stubbed my toe on the chair and **it** really hurt.

Frege

Mode of presentation (*Sinn*)
vs. reference

- The morning star/the evening star
- Mark Twain/Samuel Clemens



Worth solving?

English constraints

- Number
 - I have a car. They are blue [*they = car]
- Gender
 - My dad is shoveling snow. He's cold. [*he = snow]
- Person
 - We're watching a movie. He likes it [*he = you and I]

English exceptions

- Number
 - A: *I have a new friend.*
 - B: *What's their name?*
 - *We are a grandmother* (Margaret Thatcher)
- Gender
 - “The Nellie, a cruising yawl, swung to *her* anchor without a flutter of the sails, and was at rest.” (Heart of Darkness)
 - *It puts the lotion in the basket* (Silence of the Lambs)
- Person
 - ???

English preferences

- Recency: more recent NPs are preferred
- Grammatical role: subjects are preferred
 - Billy Bones went to the bar with Jim Hawkins. **He** called for a glass of rum.
- Repeated mention: more discourse-salient NPs are preferred
- Parallelism
 - Long John Silver went with Jim to the Old Parrot. Billy Bones went with **him** to the Old Anchor inn.
- Verb semantics
- Selectional restrictions

Verb semantics

Verb semantics

- John telephoned Bill. **He** lost the laptop

Verb semantics

- John telephoned Bill. **He** lost the laptop
- John criticized Bill. **He** lost the laptop.

Winograd challenge

<http://www.common senser reasoning.org>

Winograd challenge

- The trophy would not fit in the brown suitcase because **it** was too big
 - What does **it** refer to?

Winograd challenge

- The trophy would not fit in the brown suitcase because **it** was too ~~big~~—small
 - What does **it** refer to?

Winograd challenge

- The trophy would not fit in the brown suitcase because **it** was too ~~big~~—small
 - What does **it** refer to?
- The town councilors refused to give the demonstrators a permit because **they** feared violence.
 - Who feared violence?

Winograd challenge

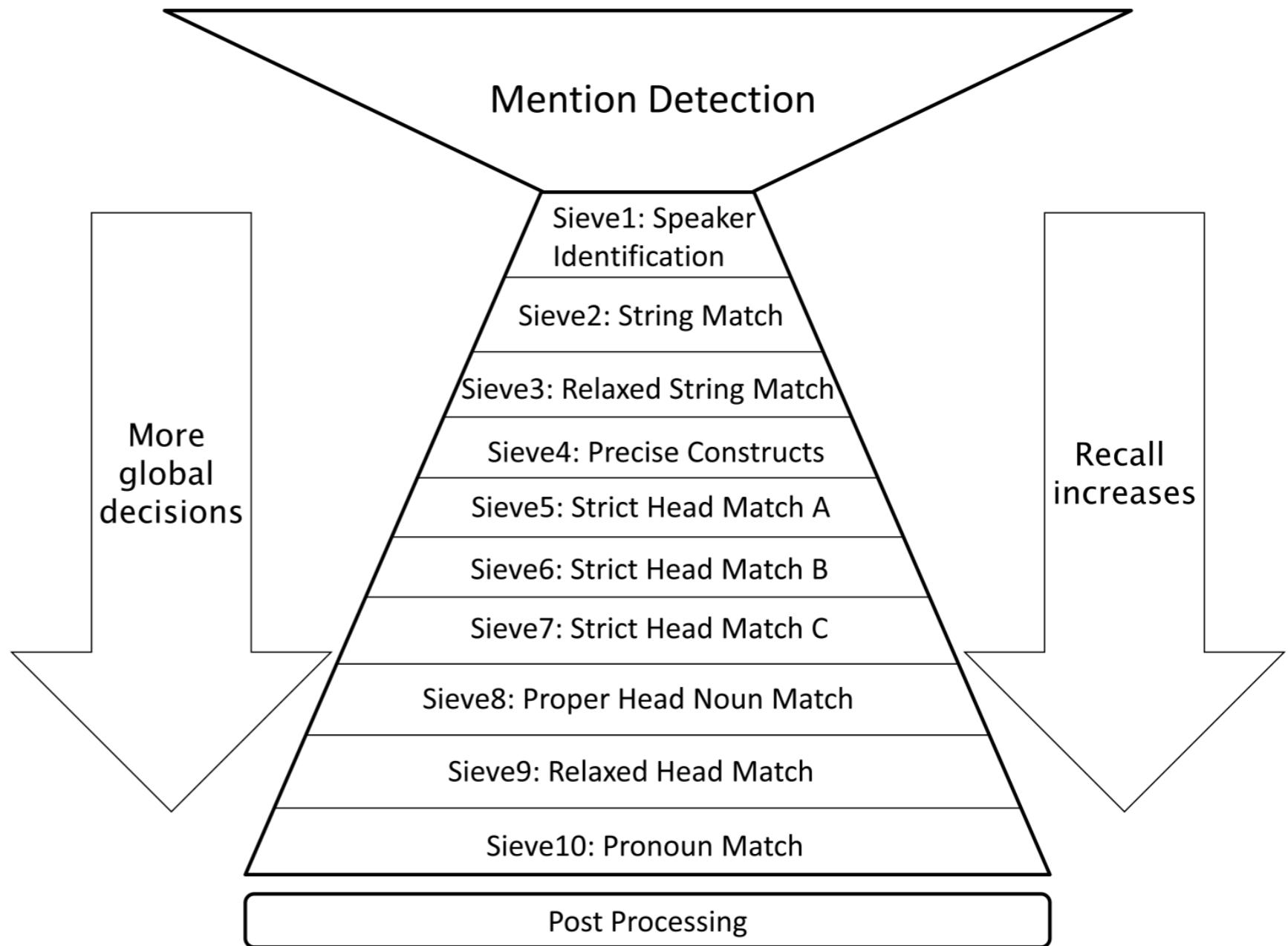
- The trophy would not fit in the brown suitcase because **it** was too ~~big~~—small
 - What does **it** refer to?
- The town councilors refused to give the demonstrators a permit because **they** feared violence.
 - Who feared violence?
- The town councilors refused to give the demonstrators a permit because **they** advocated violence.
 - Who advocated violence?

Selectational restrictions

- John parked his car in the garage after driving *it* around for hours.

Stanford “Sieve”

Sequence of pattern matching rules starting at high precision coreference links, progressing to higher recall.



Mention Detection

- All NPs, possessive pronouns, and named entity mentions are **candidate mentions**. Recall is more important than precision.
- Filters to remove candidates:
 - Remove mentions embedded within larger mentions with same headword
 - Remove numeric quantities (100 miles, 9%)
 - Remove existential there, it
 - Remove adjectival forms of nations
 - Remove 8 stop words (there, ltd., hmm)

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
"[It]⁷ is [[my]⁹ favorite]⁸," [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[**my**]⁹ favorite]⁸,” [**John**]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]⁹ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Strict Head Match B,C:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match B,C:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Proper Head Noun Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match B,C:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Proper Head Noun Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed Head Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match B,C:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Proper Head Noun Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed Head Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Pronoun Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Strict Head Match B,C:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Proper Head Noun Match:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Relaxed Head Match:

[John]¹ is [a musician]¹. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁷ is [[my]¹ favorite]⁷,” [John]¹⁰ said to [her]¹¹.

Pronoun Match:

[John]¹ is [a musician]¹. [He]¹ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁴ is [[my]¹ favorite]⁴,” [John]¹⁰ said to [her]⁵.

Post Processing:

[John]¹ is a musician. [He]¹ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁴.
“[It]⁴ is [my]¹ favorite,” [John]¹⁰ said to [her]⁵.

John is a musician.
He played a new
song. A girl was
listening to the song.
“It is my favorite,”
John said to her.

Mention Detection:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Speaker Sieve:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed String Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Precise Constructs:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match A:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Strict Head Match B,C:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Proper Head Noun Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Relaxed Head Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Pronoun Match:

[John]¹ is [a musician]². [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [[my]⁹ favorite]⁸,” [John]¹⁰ said to [her]¹¹.

Post Processing:

[John]¹ is a musician. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [my]⁹ favorite,” [John]¹⁰ said to [her]¹¹.

Final Output:

[John]¹ is a musician. [He]³ played [a new song]⁴.
[A girl]⁵ was listening to [the song]⁶.
“[It]⁷ is [my]⁹ favorite,” [John]¹⁰ said to [her]¹¹.

Classification (Refresher)

A mapping h from input data \mathbf{x} (drawn from instance space \mathcal{X}) to a label (or labels) \mathbf{y} from some enumerable output space \mathcal{Y}

\mathcal{X} = set of all documents

\mathcal{Y} = {english, mandarin, greek, ...}

\mathbf{x} = a single document

\mathbf{y} = ancient greek

Classification for Coref

A mapping h from input data $\textcolor{magenta}{x}$ (drawn from instance space \mathcal{X}) to a label (or labels) $\textcolor{magenta}{y}$ from some enumerable output space \mathcal{Y}

Positive examples = pronouns paired with closest antecedent (or coreference chain)

Negative examples = entities not in coreference chain.

Classification

For every possible antecedent y for pronoun x , we frame a binary classification: **is y coreferent with x ?** Every noun phrase is a candidate antecedent.

- I
- you
- you
- the power
- the power of the dark side
- the dark side
- Obi-Wan
- you
- your
- your father
- He
- me
- you

LUKE
I'll never join you!

VADER
If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE
He told me enough! It was you who killed him.

VADER
No. I am your father.

LUKE
No. No. That's not true!
That's impossible!

VADER
Search your feelings. You know it to be true.

LUKE
No! No! No!

Classification

For every possible antecedent y for pronoun x , we frame a binary classification: **is y coreferent with x ?** Every noun phrase is a candidate antecedent.

- I
- you
- you
- the power
- the power of the dark side
- the dark side
- Obi-Wan
- you
- your
- your father
- He
- me
- you

LUKE
I'll never join you!

VADER
If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE
He told me enough! It was you who killed him.

VADER
No. I am your father.

LUKE
No. No. That's not true!
That's impossible!

VADER
Search your feelings. You know it to be true.

LUKE
No! No! No!

Let's brainstorm a supervised classifier! (4 min activity)
In your Zoom group, come up with a list of features for this task?

Features

- John saw a beautiful 1961 Ford Falcon at the used car dealership
- He showed it to Bob.
- He bought it.

Features

- Unary features (valid of a single token)
 - token, lemma, part of speech
 - salience
- Binary features (valid of a pair of tokens)
 - number agreement (plural pronoun/plural NP)
 - compatible number (plural pronoun/??? NP)
 - gender agreement
 - compatible gender
 - sentence distance
 - Hobbs distance
 - syntax: grammatical role

Nominal coreference

- Pronominal coreference is a subset of the full coreference resolution problem because pronouns are nearly always **coreferent**.
- How would we extend the classification approach to general nominal referents?

Evaluation

- Evaluating general reference resolution (i.e., all noun phrase entities) is more complicated than straightforward accuracy/precision/recall

$$B_{precision}^3 = \frac{1}{n} \sum_i^n \frac{|Gold_i \cap System_i|}{|System_i|}$$

$$B_{recall}^3 = \frac{1}{n} \sum_i^n \frac{|Gold_i \cap System_i|}{|Gold_i|}$$

n ranges over all entities in the gold and system outputs

LUKE

I ll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

3 entities/coreference chains

LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

7 elements

{I, you, you, your, me, your, your, You}

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

7 elements
{I, you, you, your, me, your, your, You}

LUKE

I ll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

6 elements

{you, your father, you, him, I, your father}

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join **you!**

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to **your father.**

LUKE

He told me enough! It was **you** who killed **him.**

6 elements

{**you, your father, you, him, I, your father**}

VADER

No. **I am your father.**

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

2 elements
{Obi-Wan, He}

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

2 elements
{Obi-Wan, He}

LUKE

I ll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

Example system output: 4 entities

3 = {I, me, I}

8 = {you, you, you, your, you, your, your, you}

3 = {Obi-Wan, your father, your father}

2 = {He, him}

Evaluation

- More complicated than straightforward accuracy/precision/recall

$$B_{precision}^3 = \frac{1}{n} \sum_i^n \frac{|Gold_i \cap System_i|}{|System_i|}$$

$$B_{recall}^3 = \frac{1}{n} \sum_i^n \frac{|Gold_i \cap System_i|}{|Gold_i|}$$

n ranges over all entities in gold and system output



LUKE

I ll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!



LUKE

I ll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join **you!**

LUKE

I'll never join **you!**

VADER

If **you** only knew the power of the dark side. Obi-Wan never told **you** what happened to **your** father.

LUKE

He told me enough! It was **you** who killed him.

VADER

If you only knew the power of the dark side. Obi-Wan never told **you** what happened to **your father**.

LUKE

He told me enough! It was **you** who killed **him**.

VADER

No. I am **your** father.

VADER

No. **I** am **your father**.

LUKE

No. No. That's not true!
That's impossible!

LUKE

No. No. That's not true!
That's impossible!

VADER

Search **your** feelings. **You** know it to be true.

VADER

Search **your** feelings. **You** know it to be true.

LUKE

No! No! No!

LUKE

No! No! No!

LUKE
I'll never join **you!**

VADER

If **you** only knew the power of the dark side. Obi-Wan never told **you** what happened to **your** father.

LUKE
He told me enough! It was **you** who killed him.

VADER
No. I am **your** father.

LUKE
No. No. That's not true!
That's impossible!

VADER
Search **your** feelings. **You** know it to be true.

LUKE
No! No! No!

LUKE
I ll never join **you!**

VADER
If **you** only knew the power of the dark side. Obi-Wan never told **you** what happened to **your** father.

LUKE
He told **me** enough! It was **you** who killed him.

VADER
No. I am **your** father.

LUKE
No. No. That's not true!
That's impossible!

VADER
Search **your** feelings. **You** know it to be true.

LUKE
No! No! No!

LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join you!

VADER

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

LUKE

He told me enough! It was you who killed him.

VADER

No. I am your father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search your feelings. You know it to be true.

LUKE

No! No! No!

LUKE

I'll never join **you!**

VADER

↓ **you** only knew the power of the dark side. Obi-Wan never told **you** what happened to **your** father.

LUKE

He told me enough! It was **you** who killed him.

VADER

No. I am **your** father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search **your** feelings. **You** know it to be true.

LUKE

No! No! No!

LUKE

I ll never join you!

VADER

↓ **you** only knew the power of the dark side. Obi-Wan never told **you** what happened to **your** father.

LUKE

He told **me** enough! It was **you** who killed him.

VADER

No. I am **your** father.

LUKE

No. No. That's not true!
That's impossible!

VADER

Search **your** feelings. **You** know it to be true.

LUKE

No! No! No!

Evaluation

- More complicated than straightforward accuracy/precision/recall

$$B_{precision}^3 = \frac{1}{n} \sum_i^n \frac{|Gold_i \cap System_i|}{|System_i|}$$

$$B_{recall}^3 = \frac{1}{n} \sum_i^n \frac{|Gold_i \cap System_i|}{|Gold_i|}$$

n ranges over all entities in gold and system output

Hard coreference

“Between him and Darcy there was a very steady friendship, in spite of great opposition of character. Bingley was endeared to Darcy by the easiness, openness, and ductility of his temper, though no disposition could offer a greater contrast to his own, and though with his own he never appeared dissatisfied. On the strength of Darcy's regard, Bingley had the firmest reliance, and of his judgement the highest opinion. In understanding, Darcy was the superior. Bingley was by no means deficient, but Darcy was clever. He was at the same time haughty, reserved, and fastidious, and his manners, though well-bred, were not inviting. In that respect his friend had greatly the advantage. Bingley was sure of being liked wherever he appeared, Darcy was continually giving offense.”

- You cannot read *Cyril Connolly* for very long without wanting to acquire —and then developing— a relationship with the personality of *the man* himself.
[. . .] With *Connolly* there is a marked difference and the difference is that *the artist* and *the man* are so conjoined and intermingled that you cannot savour *the one* without *the other* and vice versa.

Non-identity

- Non-Identity. The two NPs point to two different DEs. Even if they share any feature, they are not ‘the same thing.’
- “President Samaranch sent **a letter** to Sydney in which he asked for information. **A similar missive** has also been received by all the candidate cities to host the Olympic Games of 1996.”

Identity

- Identity. The two NPs point to the same DE (i.e., they have the same set of attributes, as far as one can tell). They are (almost certainly) ‘the same thing.’
- “It began when **a Hasidic Jewish family** bought one of the town’s two meat-packing plants 13 years ago. First **they** brought in other Hasidic Jews, then Mexicans, Palestinians, Ukrainians.”

Identity

- Identity. The two NPs point to the same DE (i.e., they have the same set of attributes, as far as one can tell). They are (almost certainly) ‘the same thing.’
- “It began when **a Hasidic Jewish family** bought one of the town’s two meat-packing plants 13 years ago. First **they** brought in other Hasidic Jews, then Mexicans, Palestinians, Ukrainians.”

Near-identity

- A proper noun appears first, and a subsequent noun phrase refers to some aspect of the discourse entity
 - Role
 - Location
 - Organization
 - Information realization
 - Representation
 - Many (many) others

“Your father was the greatest” commented an anonymous old lady while she was shaking Alessandro’s hand —Gassman’s best known son. “I will miss **the actor**, but I will be lacking **my father** especially,” he said.

Role near-identity: A specific role or function performed by a human, animal or object, is distinguished from their other facets.

“Your father was the greatest” commented an anonymous old lady while she was shaking Alessandro’s hand —Gassman’s best known son. “I will miss **the actor**, but I will be lacking **my father** especially,” he said.

“The Jordan authorities arrested, on arriving in Iraq, an Italian pilot who violated the air embargo to this country.”

Location near-identity: The name of a location can be used to describe facets such as the physical place, the place associated with a (political) organization, the population living in that location, the ruling government, an affiliated organization, an event celebrated at that location, etc.

“The Jordan authorities arrested, on arriving in Iraq, an Italian pilot who violated the air embargo to this country.”

She hasn't seen *Gone with the Wind*, but she's read it.

Information realization near-identity: A discourse entity corresponding to an informational object (e.g., story, law, review, etc.) can be split according to the format in which the information is presented or manifested (FRBR abstraction hierarchy)

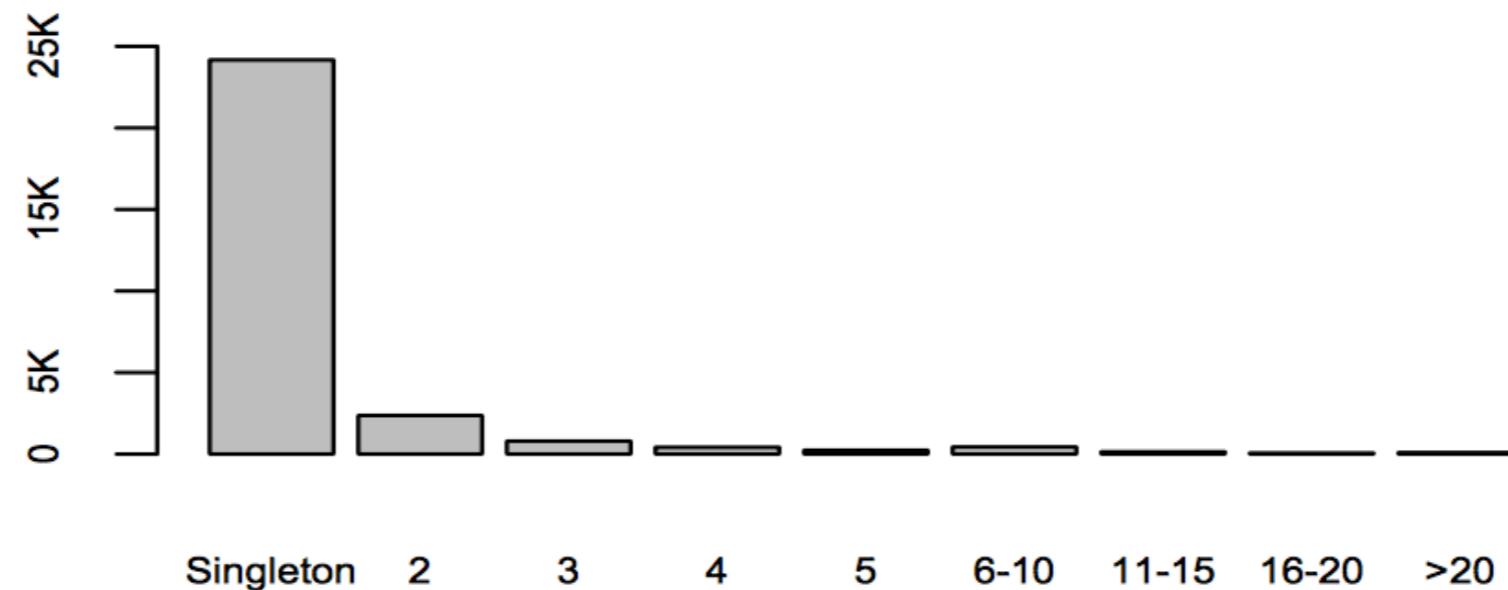
She hasn't seen *Gone with the Wind*, but she's read *it*.

Singletons

- At test time we don't have access to true mentions
[John] saw [a beautiful 1961 Ford Falcon] at [the used car dealership]. [He] showed [it] to [Bob]. [He] bought [it].

Singletons

- Most noun phrases in a discourse are not coreferent. They are **singleton** mentions.



Singletons

- We can built a classifier to predict for any noun phrase, whether it will be a part of a coreference chain or a singleton (78% accurate).

	Estimate	P-value
Type = pronoun	1.21	< 0.001
Type = proper noun	1.88	< 0.001
Animacy = inanimate	-1.36	< 0.001
Animacy = unknown	-0.38	< 0.001
Person = 1	1.05	< 0.001
Person = 2	0.13	< 0.001
Person = 3	1.62	< 0.001
Number = singular	0.61	< 0.001
Number = unknown	0.17	< 0.001
Quantifier = indefinite	-1.49	< 0.001
Quantifier = quantified	-1.23	< 0.001
Number of modifiers	-0.39	< 0.001

	Estimate	P-value
Presence of negation	-0.18	< 0.001
Presence of modality	-0.22	< 0.001
Under an attitude verb	0.03	0.01
AttitudeVerb * (Type = pronoun)	0.29	< 0.001
AttitudeVerb * (Type = proper noun)	0.14	< 0.001
Modal * (Type = pronoun)	0.12	0.04
Modal * (Type = proper noun)	0.35	< 0.001
Negation * (Type = pronoun)	1.07	< 0.001
Negation * (Type = proper noun)	0.30	< 0.001
Negation * (Quantifier = indefinite)	-0.37	< 0.001
Negation * (Quantifier = quantified)	-0.36	0.23
Negation * (Number of modifiers)	0.11	< 0.001



Surprise Exam Prep!

Knowledge powers the world these days and you just landed a job at the Wikimedia Foundation that runs all of the the different Wikipedias in multiple languages. Jimmy Wales comes to you personally with a big problem: sentences that need citations. In an era of fact checking, Jimmy is concerned that some Wikipedia content needs some kind of citation to back up its claim. However, Wikipedia is too big to check every sentence by hand, even if crowdsourcing was used (after all, not everyone donates that \$5). He wants you to develop a suite of approaches for determining whether a sentence might need a citation, starting with sentences in the English Wikipedia.

Describe at one approach for how you would develop an NLP method to detect whether a sentence needs a citation. Be sure to go into details on how you would train the method, what kind of features you would use, etc.

We'll discuss this in the breakouts rooms and then come back



Discourse

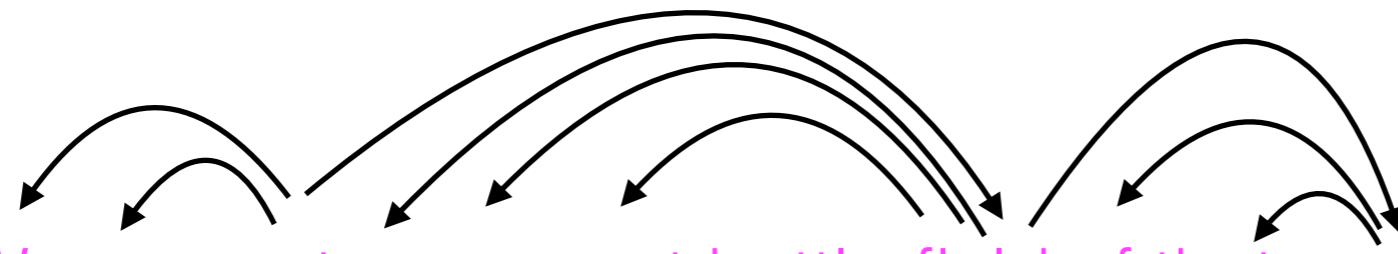
great

JJ

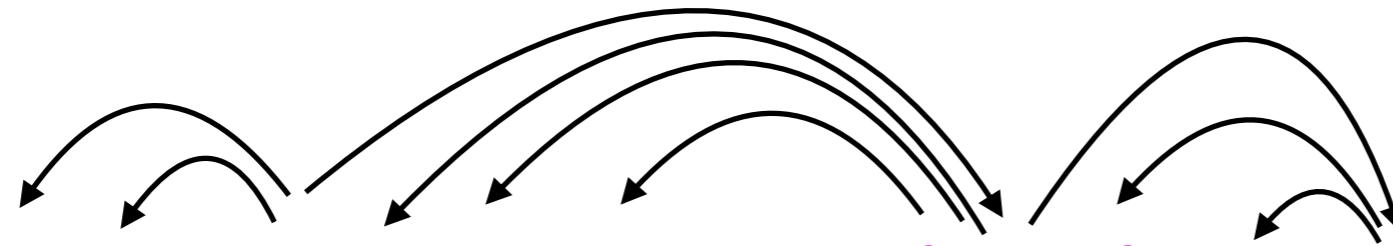
great

great

We are met on a great battle-field of that war.



We are met on a great battle-field of that war.



We are met on a great battle-field of that war.

$\lambda x. \lambda y. \text{meet}(x, y)$

$\lambda x. \text{battlefield}(x)$

$\lambda x. \text{war}(x)$

We are met on a great battle-field of that war.

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. **We are met on a great battle-field of that war.** We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we can not dedicate -- we can not consecrate -- we can not hallow -- this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us -- that from these honored dead we take increased devotion to that cause for which they gave the last full measure of devotion -- that we here highly resolve that these dead shall not have died in vain -- that this nation, under God, shall have a new birth of freedom -- and that government of the people, by the people, for the people, shall not perish from the earth.

Discourse

- Discourse covers linguistic expression **beyond the boundary of the sentence.**
 - Dialogues: the structure of turns in conversation
 - Monologues: the structure of entire passages, documents

What is the **structure** of discourse?

Discourse structure

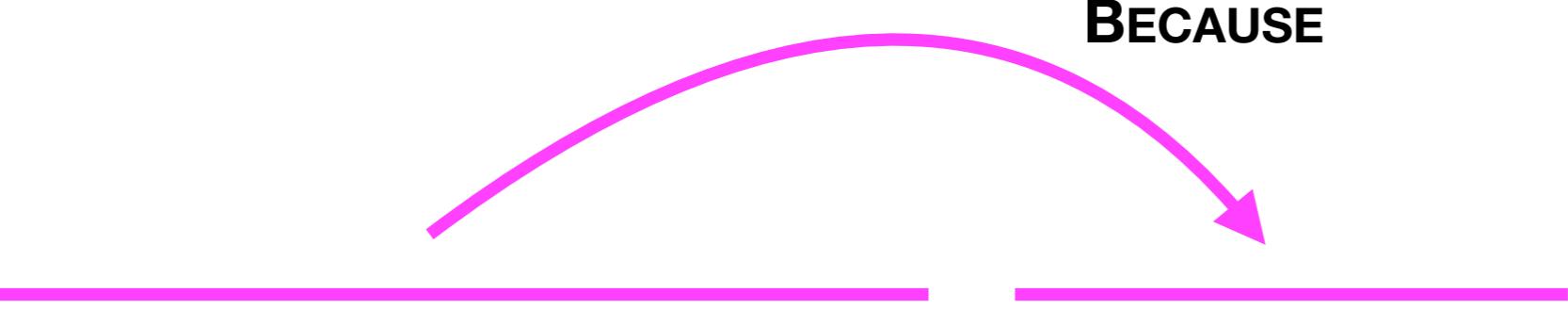
Discourse structure

- John hid Bill's car keys. He was drunk

Discourse structure

- John hid Bill's car keys. He was drunk
- John hid Bills car keys. He likes spinach.

Discourse structure

- 
- The diagram consists of two horizontal pink lines. A pink curved arrow originates from the left line and points to the right line. The word "BECAUSE" is written in black capital letters above the arrow.
- John hid Bill's car keys. He was drunk
 - John hid Bills car keys. He likes spinach.

Discourse structure

- Useful for:
 - Text summarization (Louis et al. 2010)
 - Evaluating writing (Burnstein et al. 2013)

Discourse structure

- Two representations of structure in discourse:
 - deep structure (RST)
 - shallow structure (PDTB)

RST

- Rhetorical structure theory (Mann and Thompson 1987)
- RST Corpus: 385 WSJ articles (from Penn Treebank) = 176K words.

Elementary discourse units

- Elementary discourse units: the “minimal building blocks of a discourse tree” (Carlson 2002)

Elementary discourse units

- Clauses
- Prosodic units
- Turns of talk
- Sentences

Elementary discourse units

Xerox Corp's third-quarter net income grew 6.2% on 7.3% higher revenue.

This earned mixed reviews from Wall Street analysts.

Xerox Corp's third-quarter net income grew 6.2% on 7.3% higher revenue,

which earned mixed reviews from Wall Street analysts.

Xerox Corp's third-quarter net income grew 6.2% on 7.3% higher revenue,

earning mixed reviews from Wall Street analysts.

The 6.2% growth of Xerox Corp's third quarter net revenue on 7.3% higher revenue

earned mixed reviews from Wall Street analysts.

Elementary discourse units

- Clauses
- Prosodic units
- Turns of talk
- Sentences

Elementary discourse units

- Clauses
 - Prosodic units
 - Turns of talk
 - Sentences
- 
- ```
graph LR; A["RST Discourse treebank"] --> B["• Clauses"]
```

# Clauses in RST

- Subject/object clauses (gerunds, infinitives) are not EDUs:
- Making computer smaller often means sacrificing memory.
- Insurers could see claims totaling nearly \$1 billion from the San Francisco earthquake

# Clauses in RST

- Some phrasal EDUs that begin with a strong discourse marker
  - because
  - in spite of
  - as a result of
  - according to

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.
- Surface temperatures average -60 C

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.
- Surface temperatures average -60 C
- and can dip to -123 C.

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.
- Surface temperatures average -60 C
- and can dip to -123 C.
- Only the midday sun at tropical latitudes is warm enough

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.
- Surface temperatures average -60 C
- and can dip to -123 C.
- Only the midday sun at tropical latitudes is warm enough
- to thaw the ice

With its distant orbit, Mars experiences frigid weather. Surface temperatures average -60 C and can dip to -123 C. Only the midday sun at tropical latitudes is warm enough to thaw the ice.

- With its distant orbit,
- Mars experiences frigid weather.
- Surface temperatures average -60 C
- and can dip to -123 C.
- Only the midday sun at tropical latitudes is warm enough
- to thaw the ice

EDU: ~97% IAA; auto: ~95%

# Relations

- **Adjacent** EDUs are linked with discourse relations.

# Relations

- Relations hold between a **nucleus** and **satellite**.
  - Nucleus: more salient to discourse
  - Satellite: supporting information
- Analogous to head/dependent in dependency parsing but scoped over constituents.

Mars  
experiences  
frigid weather.

Surface  
temperatures  
average -60 C

# Relations

|             |                                                              |
|-------------|--------------------------------------------------------------|
| List        | Billy was the mate and Long John was the quartermaster       |
| Elaboration | The company wouldn't say , citing competitive reasons        |
| Attribution | Analysts estimated that sales at US stores declined          |
| Same-Unit   | [two EDUS interrupted by another EDU]                        |
| Contrast    | The priest was in a bad mood but the lama was happy          |
| Background  | T is the pointer to the root of a binary tree . Initialize T |

Most common relations in RST treebank

# Relations

|             |                                                               |
|-------------|---------------------------------------------------------------|
| List        | Billy was the mate <b>and</b> Long John was the quartermaster |
| Elaboration | The company wouldn't say , <b>citing</b> competitive reasons  |
| Attribution | Analysts <b>estimated</b> that sales at US stores declined    |
| Same-Unit   | [two EDUS interrupted by another EDU]                         |
| Contrast    | The priest was in a bad mood <b>but</b> the lama was happy    |
| Background  | T is the pointer to the root of a binary tree . Initialize T  |

Most common relations in RST treebank

# Let's do a breakout exercise!

With its  
distant orbit,

Mars  
experiences  
frigid  
weather.

Surface  
temperature  
is average  
-60 C

and can dip  
to -123 C

Only the  
midday sun at  
tropical  
latitudes is  
warm enough

to thaw the  
ice

# Let's do a breakout exercise!

In your breakout group, try to link the following nuclei using the discourse relationships shared in the chat

With its distant orbit,

Mars experiences frigid weather.

Surface temperature s average -60 C

and can dip to -123 C

Only the midday sun at tropical latitudes is warm enough

to thaw the ice

With its  
distant orbit,

Mars  
experiences  
frigid  
weather.

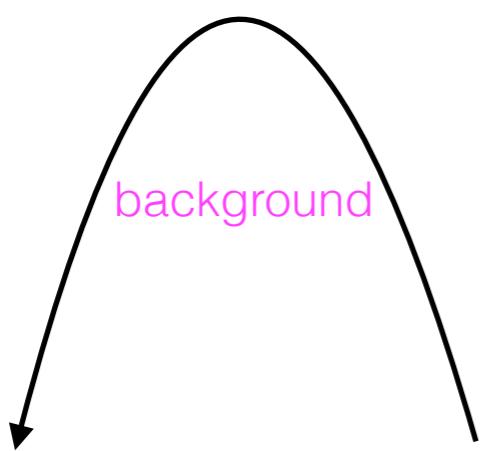
Surface  
temperature  
s average  
-60 C

and can dip  
to -123 C

Only the  
midday sun at  
tropical  
latitudes is  
warm enough

to thaw the  
ice

Arrows are directed from nucleus to satellite (undirected edges are nuclei)



With its  
distant orbit,

Mars  
experiences  
frigid  
weather.

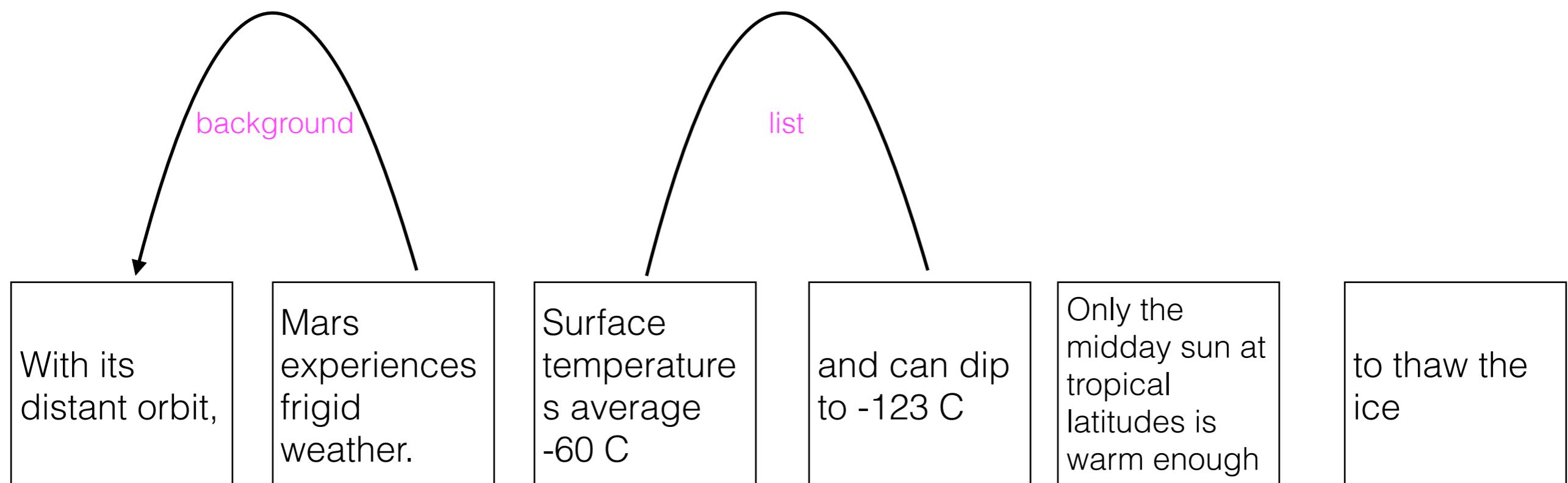
Surface  
temperatures average  
-60 C

and can dip  
to -123 C

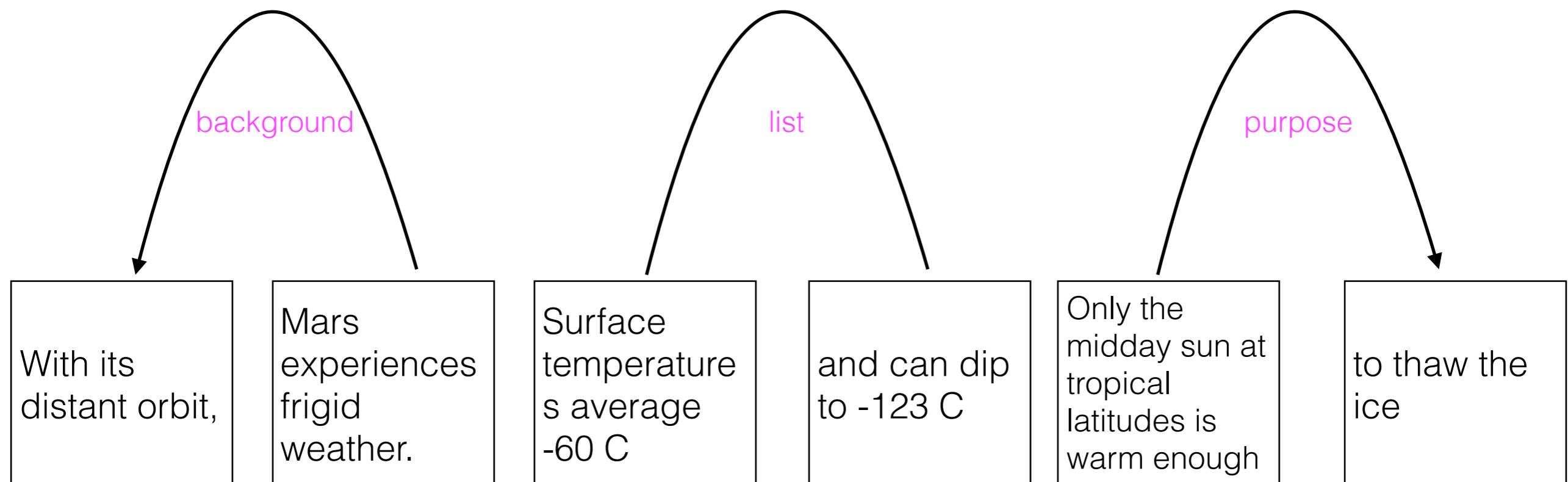
Only the  
midday sun at  
tropical  
latitudes is  
warm enough

to thaw the  
ice

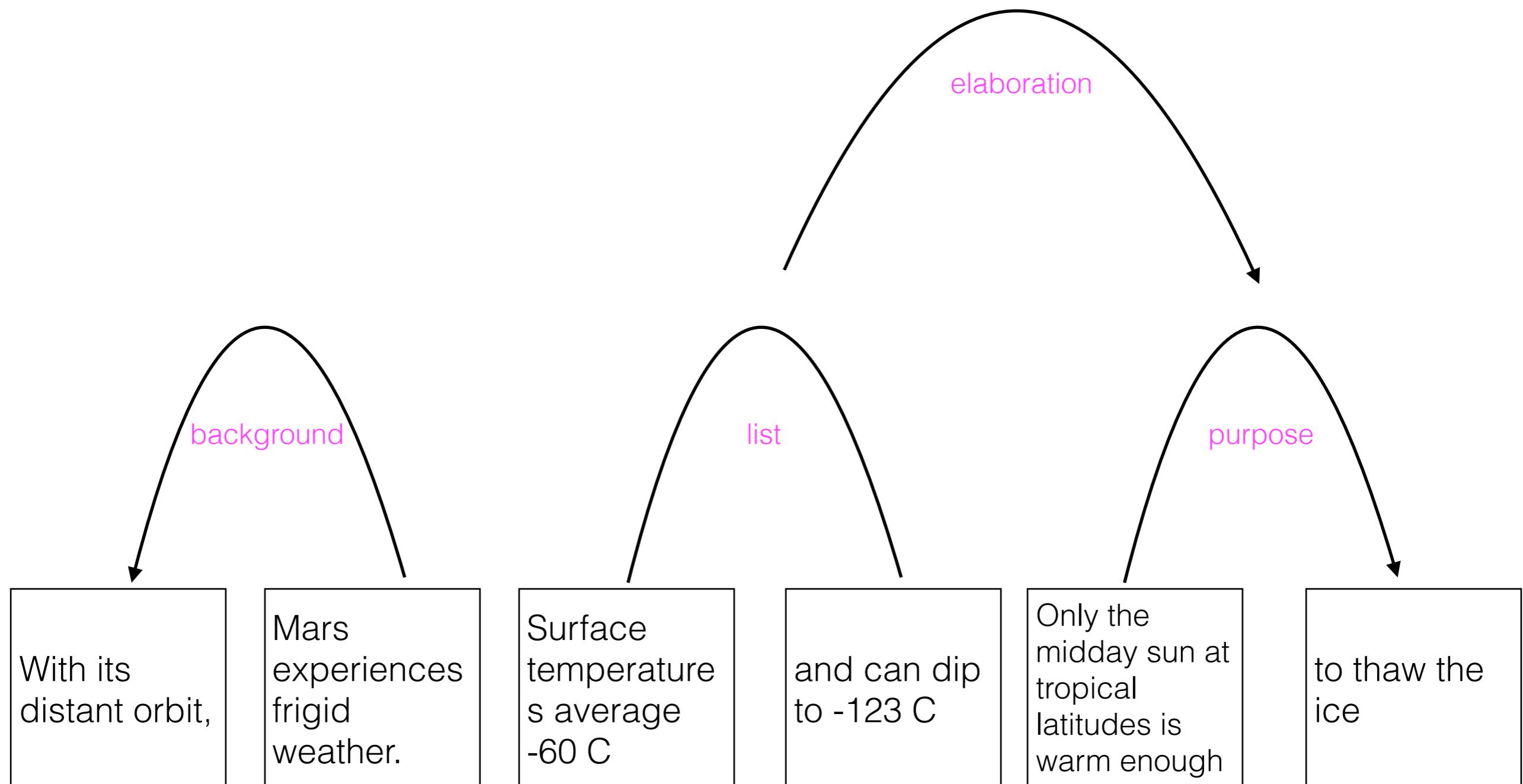
Arrows are directed from nucleus to satellite (undirected edges are nuclei)



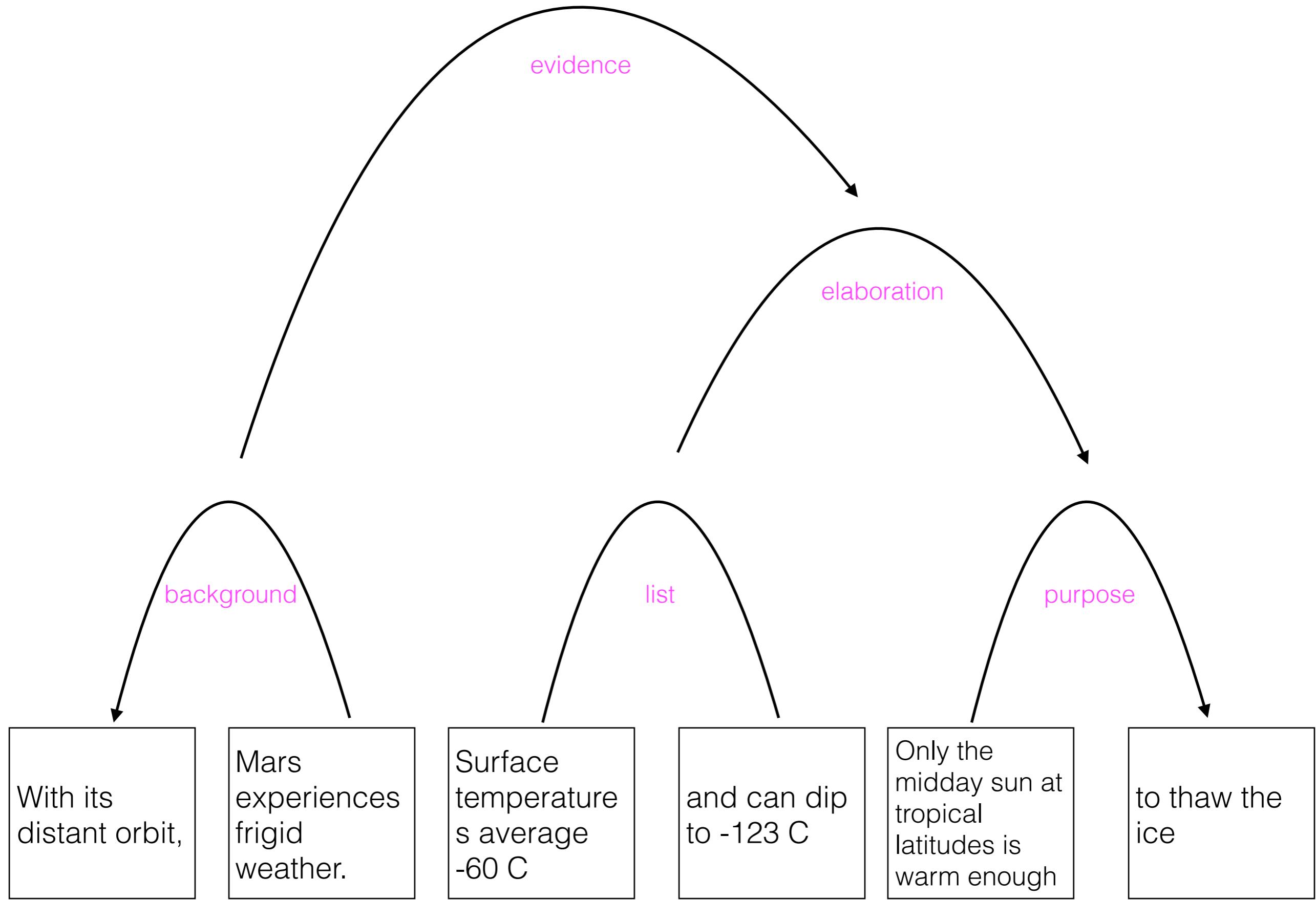
Arrows are directed from nucleus to satellite (undirected edges are nuclei)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)



Arrows are directed from nucleus to satellite (undirected edges are nuclei)

# Relations

- A given span of text will hold several (nested) relations

|                                                          |                                                                                                                                       |            |
|----------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|------------|
| With its distant orbit,                                  | Mars experiences frigid weather.                                                                                                      | background |
| With its distant orbit, Mars experiences frigid weather. | Surface temperatures average -60 C and can dip to -123 C.<br>Only the midday sun at tropical latitudes is warm enough to thaw the ice | evidence   |

# RST parsing

- Shift-reduce (Marcu 1999, Sagae 2009)
- Neural-network shift reduce (Ji and Eisenstein 2014)
- Recursive neural network (Li et al. 2014)

stack

---

With its distant orbit,

Mars experiences frigid weather.

Surface temperatures average -60 C

and can dip to -123 C

action

---

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at  $\text{stack}_1$  and subtree at  $\text{stack}_2$ : remove  $\text{stack}_1$  and  $\text{stack}_2$  from stack; create a new subtree from  $\text{stack}_1$  and  $\text{stack}_2$  and place it onto the stack

With its distant orbit,

stack

---

Mars experiences frigid weather.

Surface temperatures average -60 C

and can dip to -123 C

action

---

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at  $stack_1$  and subtree at  $stack_2$ : remove  $stack_1$  and  $stack_2$  from stack; create a new subtree from  $stack_1$  and  $stack_2$  and place it onto the stack

|                                        |        |                                                                                                                                                                                                                                                           |                          |
|----------------------------------------|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|
|                                        |        | Surface<br>temperatures<br>average -60 C                                                                                                                                                                                                                  | and can dip to<br>-123 C |
| stack                                  | action |                                                                                                                                                                                                                                                           |                          |
|                                        |        | Shift: Remove EDU from front of input buffer and push it onto stack                                                                                                                                                                                       |                          |
| Mars<br>experiences<br>frigid weather. |        | Reduce(label): assert relation between subtree at $\text{stack}_1$ and subtree at $\text{stack}_2$ : remove $\text{stack}_1$ and $\text{stack}_2$ from stack; create a new subtree from $\text{stack}_1$ and $\text{stack}_2$ and place it onto the stack |                          |
| With its distant orbit,                |        |                                                                                                                                                                                                                                                           |                          |

|               |
|---------------|
| Attribution   |
| Background    |
| Cause         |
| Comparison    |
| Condition     |
| Contrast      |
| Elaboration   |
| Enablement    |
| Evaluation    |
| Explanation   |
| Joint         |
| Manner        |
| Topic-Comment |
| Summary       |
| Temporal      |
| Topic-Change  |

Each relation generates three possible transition labels (for the position or the nucleus and satellite in the subtree)

- Attribution-NS
- Attribution-SN
- Attribution-NN



Mars  
experiences  
frigid weather.

With its distant  
orbit,

stack

---

action

---

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

background-SN:  
background relation holds between satellite on left and nucleus on right

Surface temperatures average -60 C

and can dip to -123 C

stack

---

action

---

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at  $stack_1$  and subtree at  $stack_2$ : remove  $stack_1$  and  $stack_2$  from stack; create a new subtree from  $stack_1$  and  $stack_2$  and place it onto the stack

Mars experiences frigid weather.

With its distant orbit,

background-SN:

background relation holds between satellite on left and nucleus on right

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

stack

---

action

---

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at  $stack_1$  and subtree at  $stack_2$ : remove  $stack_1$  and  $stack_2$  from stack; create a new subtree from  $stack_1$  and  $stack_2$  and place it onto the stack

With its distant orbit,

background-SN:  
background relation holds between  
satellite on left and nucleus on right

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

stack

---

action

---

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

stack

---

action

---

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

background-SN

With its distant  
orbit,

Mars  
experiences  
frigid weather.

stack

---

action

---

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

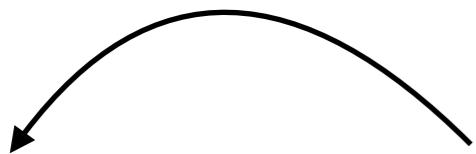
Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

background

With its distant  
orbit,

Mars  
experiences  
frigid weather.



stack

---

action

---

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

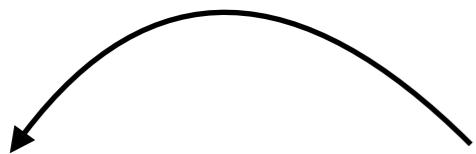
Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

background

With its distant  
orbit,

Mars  
experiences  
frigid weather.



and can dip to  
-123 C

stack

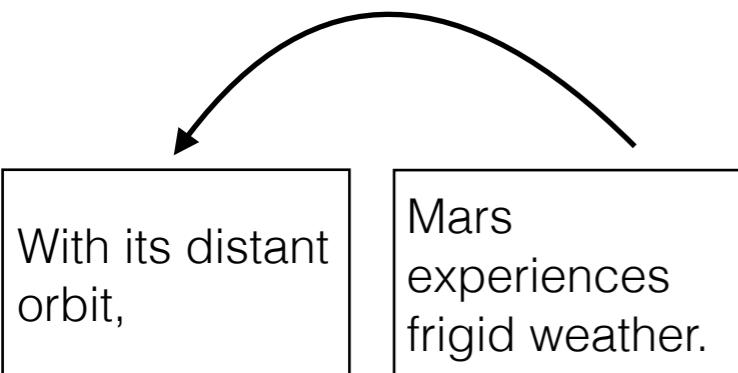
---

action

---

Surface  
temperatures  
average -60 C

background



Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at  $stack_1$  and subtree at  $stack_2$ : remove  $stack_1$  and  $stack_2$  from stack; create a new subtree from  $stack_1$  and  $stack_2$  and place it onto the stack

and can dip to  
-123 C

stack

---

action

---

Surface  
temperatures  
average -60 C

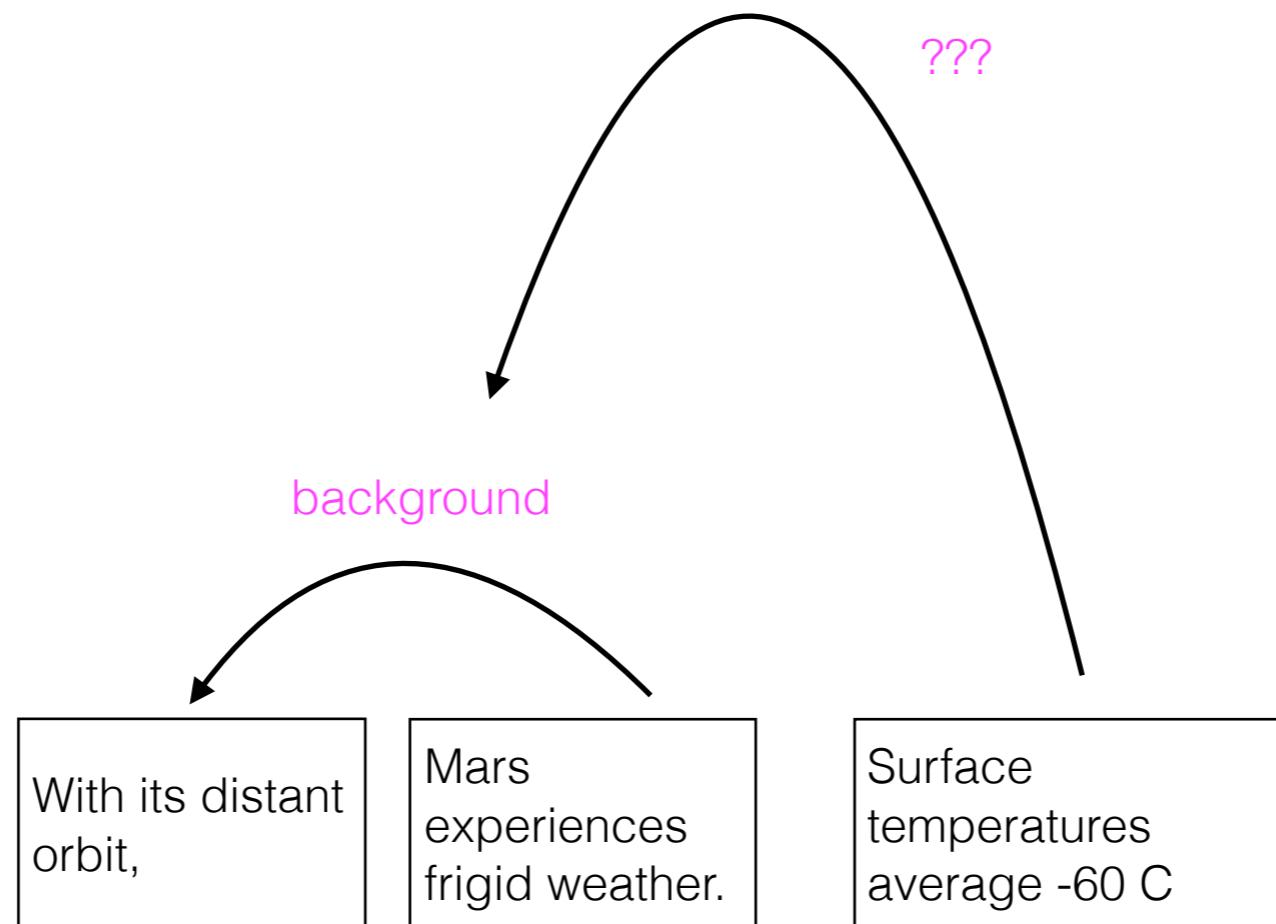
background

With its distant  
orbit,

Mars  
experiences  
frigid weather.

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack



and can dip to  
-123 C

stack

---

action

---

Surface  
temperatures  
average -60 C

background

With its distant  
orbit,

Mars  
experiences  
frigid weather.

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

stack

---

and can dip to  
-123 C

Surface  
temperatures  
average -60 C

background

With its distant  
orbit,

Mars  
experiences  
frigid weather.

action

---

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

## list-NN:

list relation holds between nucleus on left  
and nucleus on right

stack

---

and can dip to  
-123 C

Surface  
temperatures  
average -60 C

background

With its distant  
orbit,

Mars  
experiences  
frigid weather.

action

---

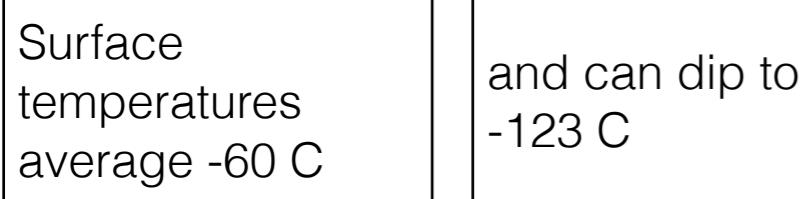
Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

stack

---

list



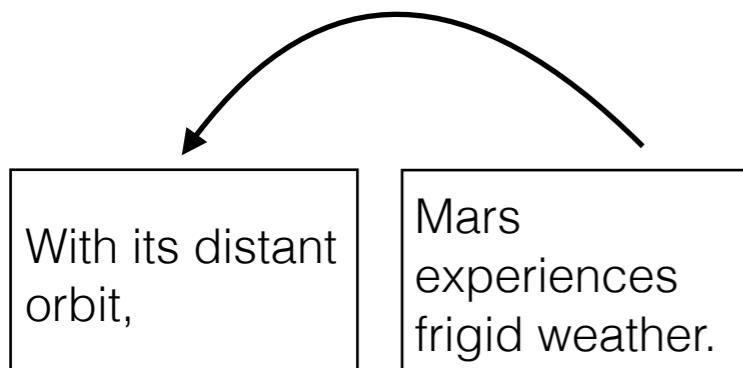
action

---

Shift: Remove EDU from front of input buffer and push it onto stack

Reduce(label): assert relation between subtree at  $stack_1$  and subtree at  $stack_2$ : remove  $stack_1$  and  $stack_2$  from stack; create a new subtree from  $stack_1$  and  $stack_2$  and place it onto the stack

background



Continue in this way until the buffer is empty and stack has a single tree

stack

---

list

Surface  
temperatures  
average -60 C

and can dip to  
-123 C

action

---

Shift: Remove EDU from  
front of input buffer and  
push it onto stack

Reduce(label): assert  
relation between subtree at  
 $stack_1$  and subtree at  $stack_2$ :  
remove  $stack_1$  and  $stack_2$   
from stack; create a new  
subtree from  $stack_1$  and  
 $stack_2$  and place it onto the  
stack

background

With its distant  
orbit,

Mars  
experiences  
frigid weather.

- RST builds a rhetorical tree over the entire span of a sentence = **deep** discourse parsing.
- SOTA for getting the labeled relations + heads right (assuming gold EDU): **61.8 F**

- The Penn Discourse Treebank encodes a different approach: **shallow** discourse parsing.
- Rather than focusing on the tree spanning the entire sentence, it focuses on recognizing the structure around **discourse connectives**.

# Discourse connectives

- Discourse connectives are those words and phrases that link together the content or purpose of adjacent spans
  - John eats porridge for breakfast, **while** Mary eats muesli.
  - Eat your porridge. **Otherwise** you're not going to football practice.
  - You've eaten your porridge every day this week. **As a result** I'm going to give you the iPod I promised you.

# Discourse connectives

- Discourse connectives are binary predicates that require two arguments

| connective  | sense               | count |
|-------------|---------------------|-------|
| and         | conjunction         | 3000  |
| but         | contrast            | 3308  |
| if          | hypothetical        | 1223  |
| when        | synchrony           | 989   |
| while       | contrast, synchrony | 781   |
| so          | result              | 263   |
| in addition | conjunction         | 165   |

# Discourse connectives

- Discourse connectives are binary predicates that require two arguments
  - [John eats porridge for breakfast] **while** [Mary eats muesli]
  - [Eat your porridge.] **Otherwise** [you're not going to football practice.]
  - [You've eaten your porridge every day this week] **As a result** [I'm going to give you the iPod I promised you.]

# Discourse connectives

- Those arguments can be in the same sentence as the connective, adjacent sentences, or non-adjacent sentences.
- The federal government suspended sales of U.S. savings bonds because *Congress hasn't lifted the ceiling on government debt.*
- Because *it operates on a fiscal year*, Bear Stearns's yearly filings are available much earlier than those of other firms.
- *Most oil companies*, when *they set exploration and production budgets for this year*, forecast revenue of \$15 for each barrel of crude produced.

# Discourse connectives

- Each discourse connective also is also tagged with its **sense**. Most words almost always occur with one sense.

|             |                                                                                                                     |
|-------------|---------------------------------------------------------------------------------------------------------------------|
| temporal    | the situations described in the arguments are related temporally (before, after, while, <b>since</b> )              |
| contingency | one argument causally influences the other (because, if, so, <b>since</b> )                                         |
| comparison  | highlights prominent differences between arguments (but, however, although)                                         |
| expansion   | expand discourse and move exposition forward (in particular, specifically, in other words); conjunction; list (and) |

# Discourse connectives

| coarse sense | fine sense                                                            |
|--------------|-----------------------------------------------------------------------|
| temporal     | Asynchronous, synchrony                                               |
| contingency  | Cause, pragmatic cause, condition, pragmatic condition                |
| comparison   | Contrast, pragmatic contrast, concession, pragmatic concession        |
| expansion    | Conjunction, instantiation, restatement, alternative, exception, list |

# Discourse connectives

- Annotation/prediction task: for all explicit discourse connectives in a text:
  - What is the discourse sense of the connective?
  - What are the spans in the text that correspond to its arguments?

# Discourse connectives

I came.

I saw.

I conquered.

# Discourse connectives

I came.

and I saw.

I conquered.

# Discourse connectives

I came.

and I saw.

and I conquered.

# Discourse connectives

The \$6 billion that some 40 companies are looking to raise in the year ending March 31 compares with only \$2.7 billion raised on the capital market in the previous fiscal year.

In fiscal 1984 before Mr. Gandhi came to power, only \$810 million was raised.

# Discourse connectives

The \$6 billion that some 40 companies are looking to raise in the year ending March 31 compares with only \$2.7 billion raised on the capital market in the previous fiscal year.

In contrast, In fiscal 1984 before Mr. Gandhi came to power, only \$810 million was raised.

# Discourse connectives

The gruff financier recently started socializing in upper-class circles.

Last year he attended a New York gala where his daughter made her debut.

# Discourse connectives

The gruff financier recently started socializing in upper-class circles.

For example, Last year he attended a New York gala where his daughter made her debut.

# Discourse connectives

John was tired. As a result he left early.

John was tired. He left early.

# Discourse connectives

What is the implicit relationship between sentences with no explicit discourse connector?

# SOTA

- CoNLL 2015 shared task on shallow discourse parsing using PDTB
- End-to-end: input = just text; output = PDTB-style discourse annotations (implicit and explicit connectors, their arguments and senses).

<http://www.cs.brandeis.edu/~clp/conll15st/results.html>

# Pipeline

- Connective classifier: identifying explicit discourse connectives
  - Argument extractor
  - Sense classifier
- 
- For all adjacent sentences, classify its sense into its implicit relation (including NoRel).

|                               | on PDTB Section 23 |       |              |              |   |           |
|-------------------------------|--------------------|-------|--------------|--------------|---|-----------|
|                               | our parser         |       |              | Lin's parser |   |           |
|                               | P                  | R     | $F_1$ (%)    | P            | R | $F_1$ (%) |
| Explicit connective           | 94.83              | 93.49 | <b>94.16</b> | -            | - | 93.62     |
| Explicit Arg1 extraction      | 51.05              | 50.33 | <b>50.68</b> | -            | - | 47.68     |
| Explicit Arg2 extraction      | 77.89              | 76.79 | <b>77.33</b> | -            | - | 70.27     |
| Explicit Both extraction      | 45.54              | 44.90 | <b>45.22</b> | -            | - | 40.37     |
| Explicit only sense           | 35.52              | 34.69 | 34.93        | -            | - | -         |
| Non-Explicit Arg1 extraction  | 64.83              | 69.50 | 67.08        | -            | - | -         |
| Non-Explicit Arg2 extraction  | 66.02              | 70.78 | 68.32        | -            | - | -         |
| Non-Explicit Both extraction  | 51.20              | 54.89 | 52.98        | -            | - | -         |
| Non-Explicit only sense       | 53.18              | 10.45 | 9.06         | -            | - | -         |
| All Arg1 extraction           | 59.20              | 61.03 | 60.10        | -            | - | -         |
| All Arg2 extraction           | 71.43              | 73.64 | 72.52        | -            | - | -         |
| All Both extraction           | 48.62              | 50.13 | 49.36        | -            | - | -         |
| Sense (Explicit+Non-Explicit) | 31.44              | 30.42 | 29.83        | -            | - | -         |
| Overall Parser                | 29.27              | 30.08 | 29.72        | -            | - | 20.64     |

|                                         | WSJ Test Set   |                |      |      |                |
|-----------------------------------------|----------------|----------------|------|------|----------------|
|                                         | 2015           |                | 2016 |      | OPT            |
|                                         | F <sub>1</sub> | F <sub>1</sub> | P    | R    | F <sub>1</sub> |
| <b>Explicit Connectives</b>             | 94.8           | <b>98.9</b>    | 96.4 | 92.5 | 94.4           |
| <b>Explicit Arg1 Extraction</b>         | 50.7           | <b>53.8</b>    | 53.1 | 50.9 | 52.0           |
| <b>Explicit Arg2 Extraction</b>         | <b>77.4</b>    | 76.7           | 74.1 | 71.1 | 72.6           |
| <b>Explicit Both Extraction</b>         | 45.2           | <b>45.3</b>    | 44.9 | 43.0 | 43.9           |
| <b>Explicit Sense Micro-Average</b>     |                |                | 38.6 | 40.2 | 39.4           |
| <b>Non-Explicit Arg1 Extraction</b>     | 67.2           | 69.9           | 72.0 | 68.0 | <b>69.9</b>    |
| <b>Non-Explicit Arg2 Extraction</b>     | 68.4           | 71.5           | 73.5 | 69.5 | <b>71.5</b>    |
| <b>Non-Explicit Both Extraction</b>     | 53.1           | 53.5           | 55.0 | 52.0 | <b>53.5</b>    |
| <b>Non-Explicit Sense Micro-Average</b> |                |                | 17.5 | 18.6 | 18.0           |
| <b>All Both Extraction</b>              | 49.4           | <b>49.6</b>    | 50.2 | 47.8 | 48.9           |
| <b>Overall Parser Performance</b>       | 29.7           | <b>30.7</b>    | 27.5 | 28.9 | 28.2           |

Open et al. (2016), “OPT: Oslo–Potsdam–Teesside: Pipelining Rules, Rankers, and Classifier Ensembles for Shallow Discourse Parsing”

# Data

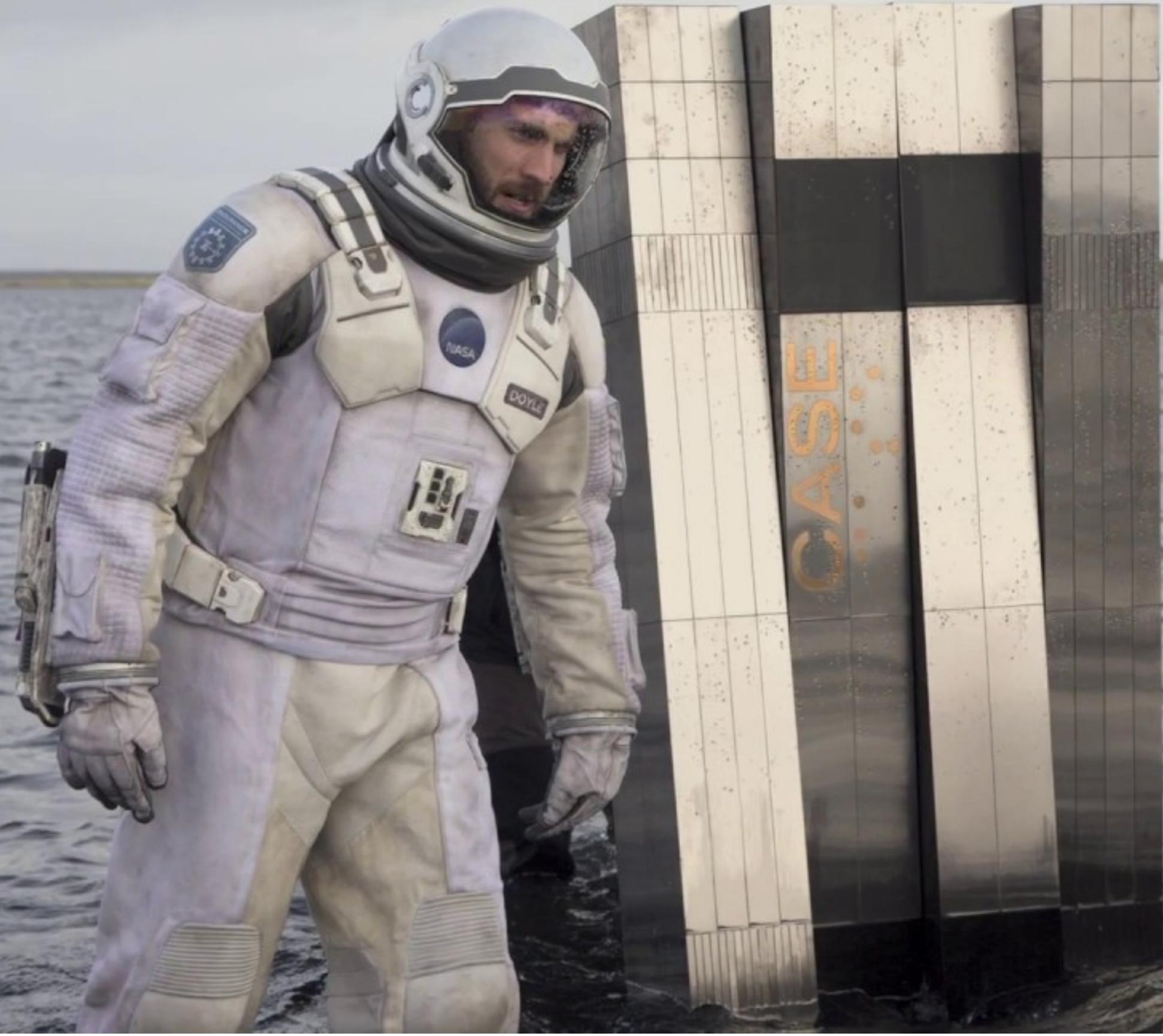
- RST Treebank [UCB Library]  
<http://digitalassets.lib.berkeley.edu/UCBonly/ldc/2002T07/LDC2002T07.tgz>
- Penn Discourse Treebank 2.0 [UCB Library]  
<http://digitalassets.lib.berkeley.edu/UCBonly/ldc/2008T05/>

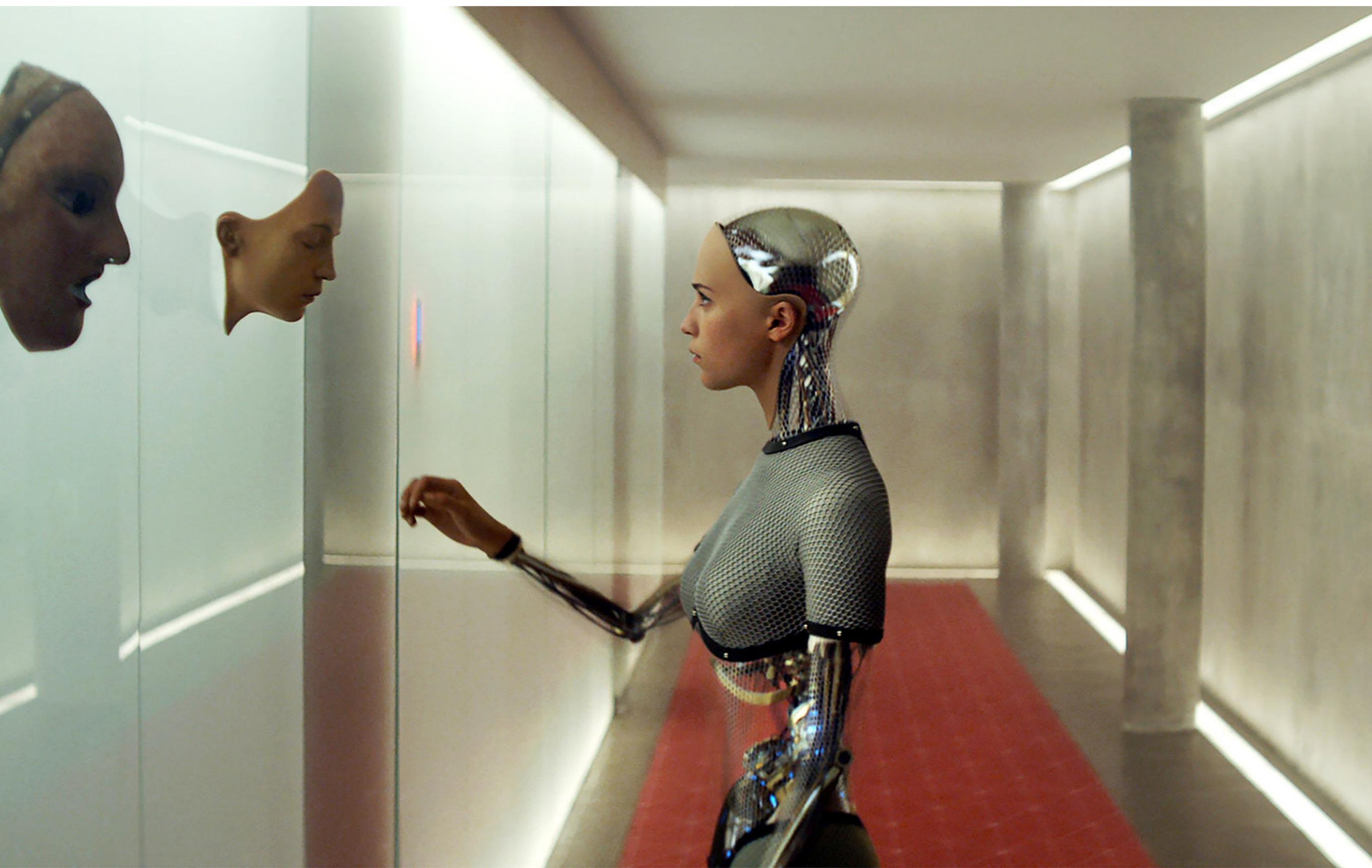
I WILL NOT COPY AGAIN I WILL NOT

I WILL NOT COPY AGAIN I



# Dialogue

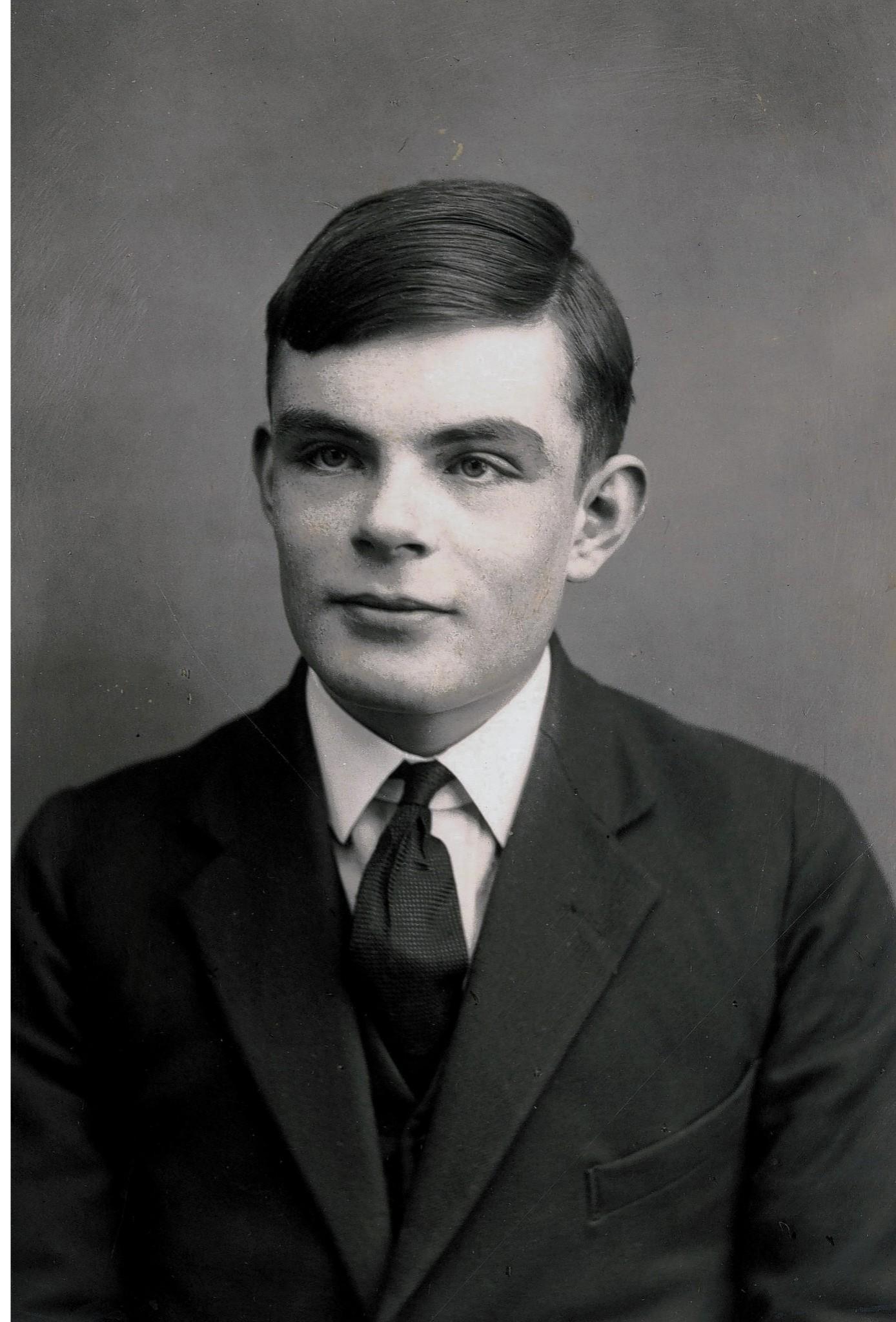




# Turing test

Distinguishing human vs.  
computer only through  
written language

Turing 1950



# Conversational bots

- Task-oriented dialog agents
  - Solicit information from users in order to complete some task (e.g., booking airline tickets, routing to customer service). “Where would you like to go?”
- Conversational chatbots
  - Systems designed for open-ended conversations with users. “How are you feeling today?”

# Conversational bots

- Task-oriented dialog agents
  - Solicit information from users in order to complete some task (e.g., booking airline tickets, routing to customer service). “Where would you like to go?”
- Conversational chatbots
  - Systems designed for open-ended conversations with users. “How are you feeling today?”

Task-oriented agents are generally *much* easier to build.

# ELIZA

- Early conversational agent (Weizenbaum 1964)
- Imitated Rogerian psychotherapist



I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

Why do you say too much work?

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

Why do you say too much work?

I am working a lot.

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

Why do you say too much work?

I am working a lot.

Earlier you said you felt unhappy?

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

Why do you say too much work?

I am working a lot.

Earlier you said you felt unhappy?

Yes.

I am the psychotherapist. Please, describe your problems.  
Each time you are finished talking, type RET twice.

I am unhappy.

Are you unhappy often?

Yes.

Can you elaborate on that?

Too much work.

Why do you say too much work?

I am working a lot.

Earlier you said you felt unhappy?

Yes.

I would appreciate it if you would continue.

# Rule-based systems

# Rule-based systems

- 0 YOU 0 ME [pattern]

# Rule-based systems

- 0 YOU 0 ME [pattern]
- WHAT MAKES YOU THINK I 3 YOU [transform]

# Rule-based systems

- 0 YOU 0 ME [pattern]
- WHAT MAKES YOU THINK I 3 YOU [transform]

# Rule-based systems

- 0 YOU 0 ME [pattern]
- WHAT MAKES YOU THINK I 3 YOU [transform]
- I think you **hate** me

# Rule-based systems

- O YOU O ME [pattern]
- WHAT MAKES YOU THINK I 3 YOU [transform]
- I think you **hate** me
- WHAT MAKES YOU THINK I **HATE** YOU

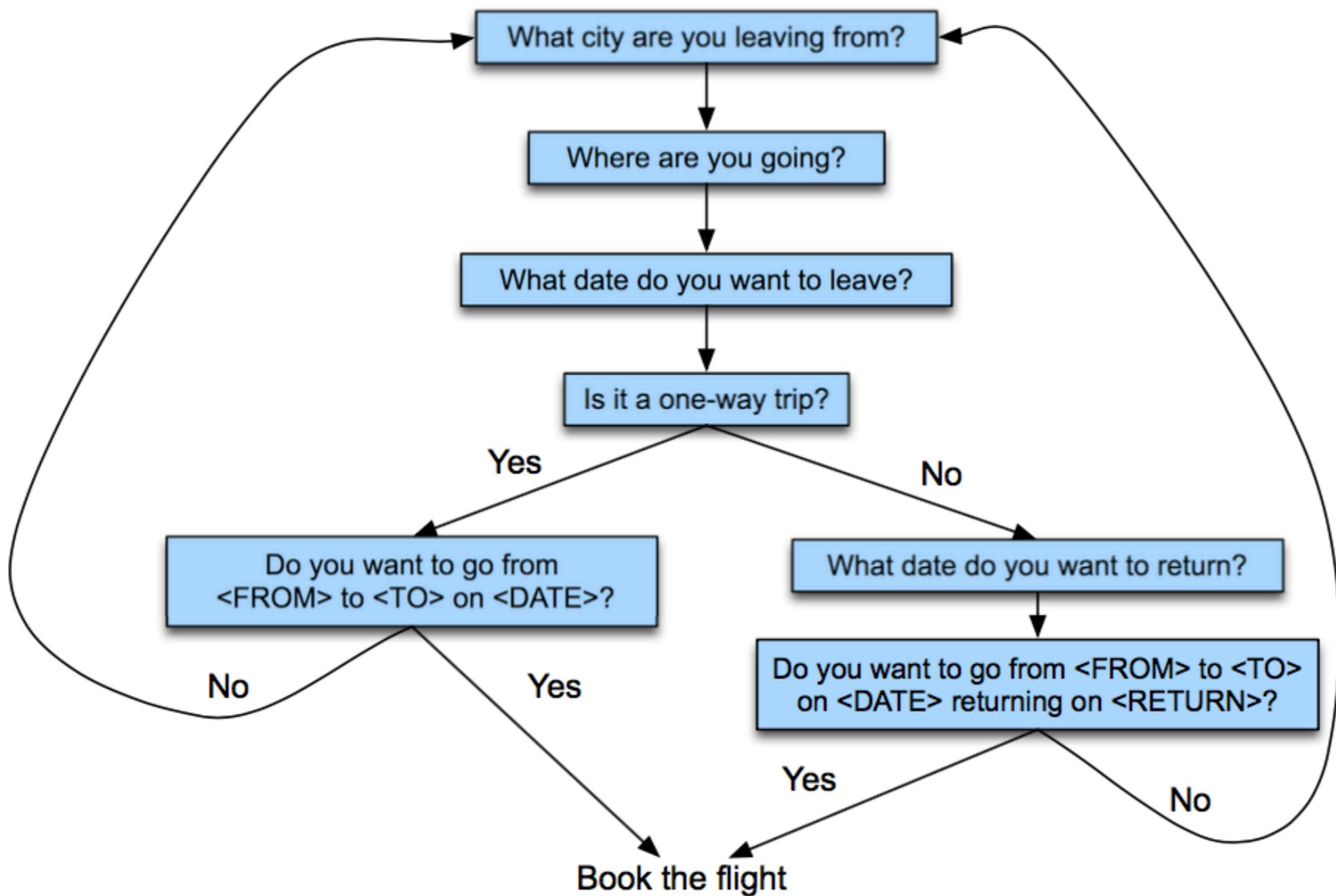
# Frame-based dialog agents

- Most modern commercial systems are based in GUS system for travel (Bobrow et al. 1977)

# Frame-based dialog agents

- **Frames** are a collection of **slots**, and each slot can take one out of a set of **values**.

| slot             | type |
|------------------|------|
| origin city      | city |
| destination city | city |
| departure time   | time |
| departure date   | date |
| arrival time     | time |
| arrival date     | date |



| slot             | type | value |
|------------------|------|-------|
| origin city      | city |       |
| destination city | city |       |
| departure time   | time |       |
| departure date   | date |       |
| arrival time     | time |       |
| arrival date     | date |       |

I'd like to book a flight to Chicago

| slot             | type | value |
|------------------|------|-------|
| origin city      | city |       |
| destination city | city |       |
| departure time   | time |       |
| departure date   | date |       |
| arrival time     | time |       |
| arrival date     | date |       |

I'd like to book a flight to Chicago

| slot             | type | value |
|------------------|------|-------|
| origin city      | city |       |
| destination city | city |       |
| departure time   | time |       |
| departure date   | date |       |
| arrival time     | time |       |
| arrival date     | date |       |

Given the available slots and the dialogue history, which slot (if any) does the turn fill?

| slot             | type | value         |
|------------------|------|---------------|
| origin city      | city | San Francisco |
| destination city | city |               |
| departure time   | time |               |
| departure date   | date |               |
| arrival time     | time |               |
| arrival date     | date |               |

| slot             | type | value         |
|------------------|------|---------------|
| origin city      | city | San Francisco |
| destination city | city |               |
| departure time   | time |               |
| departure date   | date |               |
| arrival time     | time |               |
| arrival date     | date |               |

Where from?

# San Francisco

| slot             | type | value         |
|------------------|------|---------------|
| origin city      | city | San Francisco |
| destination city | city |               |
| departure time   | time |               |
| departure date   | date |               |
| arrival time     | time |               |
| arrival date     | date |               |

Where from?

| slot             | type | value         |
|------------------|------|---------------|
| origin city      | city | San Francisco |
| destination city | city | Chicago       |
| departure time   | time |               |
| departure date   | date |               |
| arrival time     | time |               |
| arrival date     | date |               |

| slot             | type | value         |
|------------------|------|---------------|
| origin city      | city | San Francisco |
| destination city | city | Chicago       |
| departure time   | time |               |
| departure date   | date |               |
| arrival time     | time |               |
| arrival date     | date |               |

What time are you looking to leave?

| slot             | type | value         |
|------------------|------|---------------|
| origin city      | city | San Francisco |
| destination city | city | Chicago       |
| departure time   | time | 8:10          |
| departure date   | date | 11/14/17      |
| arrival time     | time | 5:10          |
| arrival date     | date | 11/14/17      |

# Tasks

- Domain classification (flights, schedule meeting, etc.)
- Intent determination (in flight domain → book a flight)
- Slot filling (the book a flight frame, find the values that fill those roles)

# Dialog agents

- Is there a notion of **frame** that can be used to structure your conversations?

| slot             | type |
|------------------|------|
| origin city      | city |
| destination city | city |
| departure time   | time |
| departure date   | date |
| arrival time     | time |
| arrival date     | date |

# Evaluation: user satisfaction

|                          |                                                             |
|--------------------------|-------------------------------------------------------------|
| <b>TTS Performance</b>   | Was the system easy to understand ?                         |
| <b>ASR Performance</b>   | Did the system understand what you said?                    |
| <b>Task Ease</b>         | Was it easy to find the message/flight/train you wanted?    |
| <b>Interaction Pace</b>  | Was the pace of interaction with the system appropriate?    |
| <b>User Expertise</b>    | Did you know what you could say at each point?              |
| <b>System Response</b>   | How often was the system sluggish and slow to reply to you? |
| <b>Expected Behavior</b> | Did the system work the way you expected it to?             |
| <b>Future Use</b>        | Do you think you'd use the system in the future?            |

**Figure 29.14** User satisfaction survey, adapted from [Walker et al. \(2001\)](#).

# Conversational Agents





say to cleverbot...



**think about it**

**think for me**

**thoughts so far**

<http://www.cleverbot.com>

## DoNotPay

Welcome. I am a bot to help you appeal your parking ticket for incorrect signage in San Francisco. Firstly, what is your ticket number?

Type your answer here ...

Press enter ↵ to send

# Dialogue as IR

- For a given turn, find the turn with the highest match in a dataset
- Return the **following** turn.

$$\cos(x, y) = \frac{\sum_{i=1}^F x_i y_i}{\sqrt{\sum_{i=1}^F x_i^2} \sqrt{\sum_{i=1}^F y_i^2}}$$



...

I'm pretty sure that's not true

**LUKE**

I'll never join you!

**VADER**

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

**LUKE**

He told me enough! It was you who killed him.

**VADER**

No. I am your father.

...

I'm pretty sure that's not true



**LUKE**

No. No. That's not true!  
That's impossible!

**VADER**

Search your feelings. You know it to be true.

**LUKE**

No! No! No!

**LUKE**

I'll never join you!

**VADER**

If you only knew the power of the dark side. Obi-Wan never told you what happened to your father.

**LUKE**

He told me enough! It was you who killed him.

**VADER**

No. I am your father.

...

I'm pretty sure that's not true



Search your feelings. You know it to be true

**LUKE**

No. No. That's not true!  
That's impossible!

**VADER**

Search your feelings. You know it to be true.

**LUKE**

No! No! No!

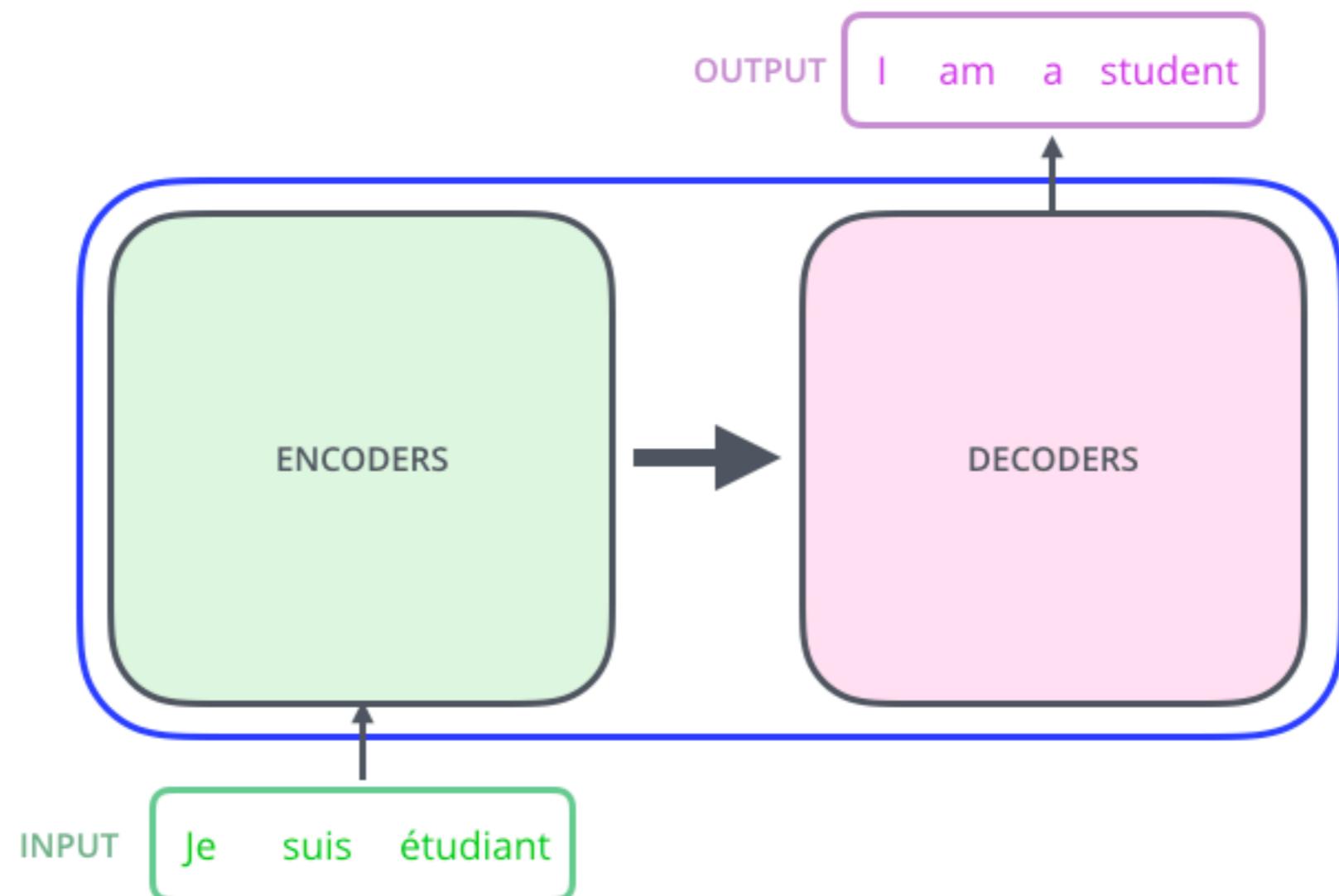
# Neural models

- Basic idea: transform a user dialogue turn into a response by the system.

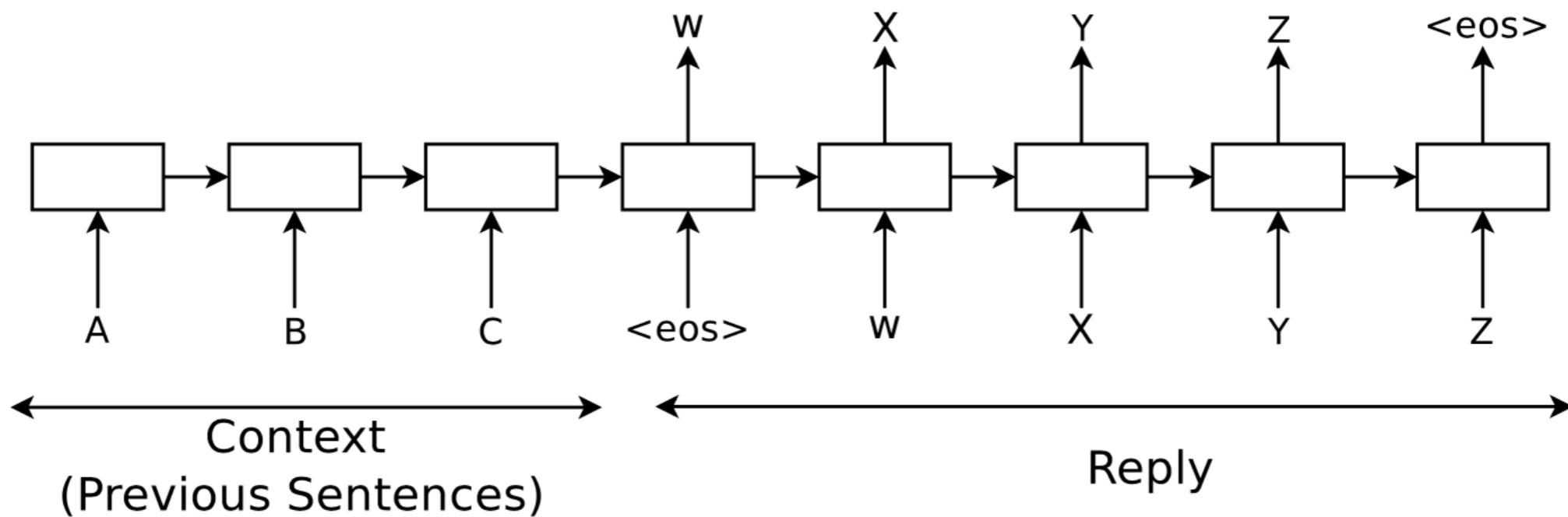
# Encoder-decoder framework

- Language modeling: predict a word given its left context
- Conversation: predict a word given its left context **and the dialogue context.**
- Machine translation: predict a word given its left context **and the full text of the source.**
- Basic idea: **encode** some context into a fixed vector; and then **decode** a new sentence from that embedding.

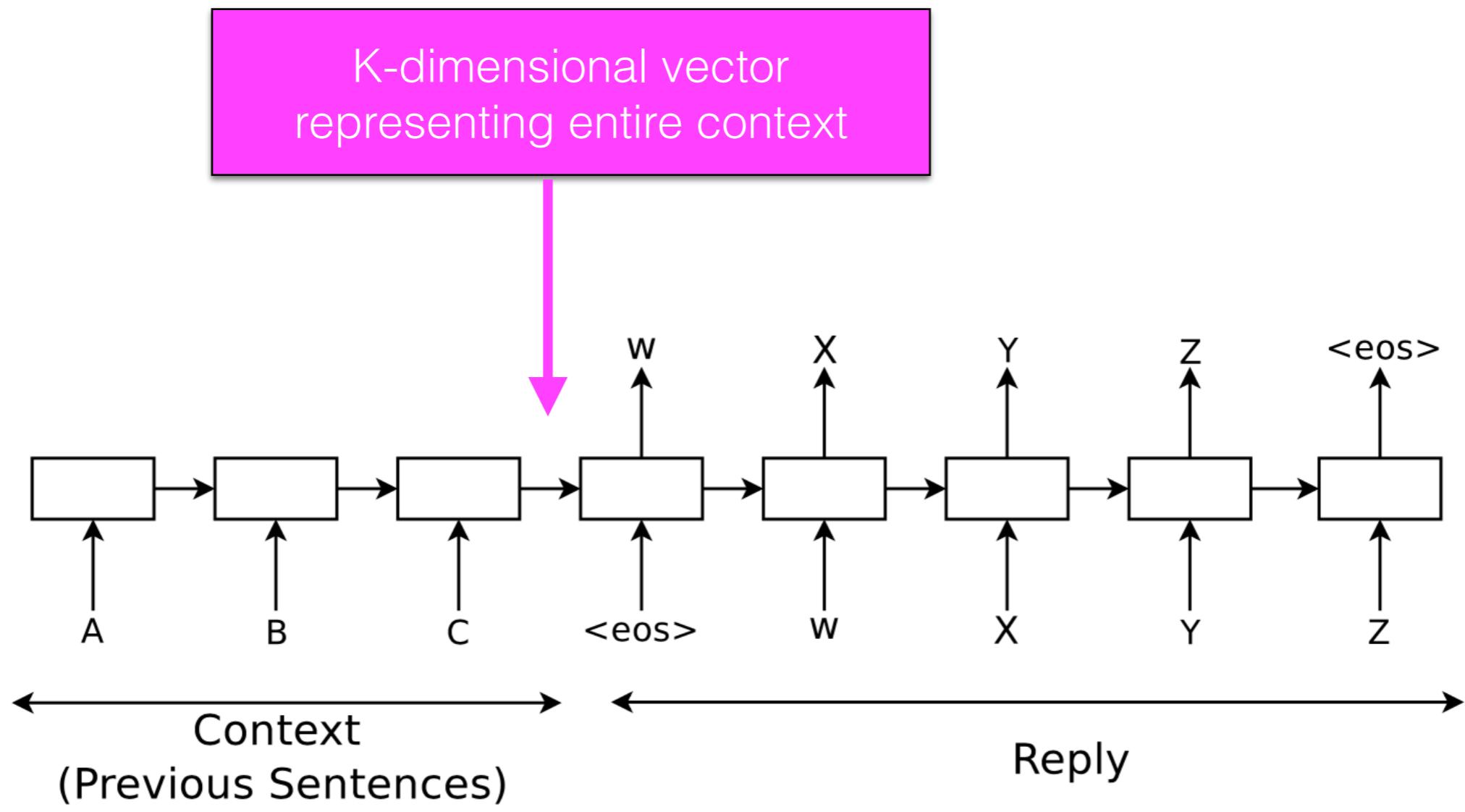
Reminder: we talked about **transformer**-based Encoder-Decoder for translation last week



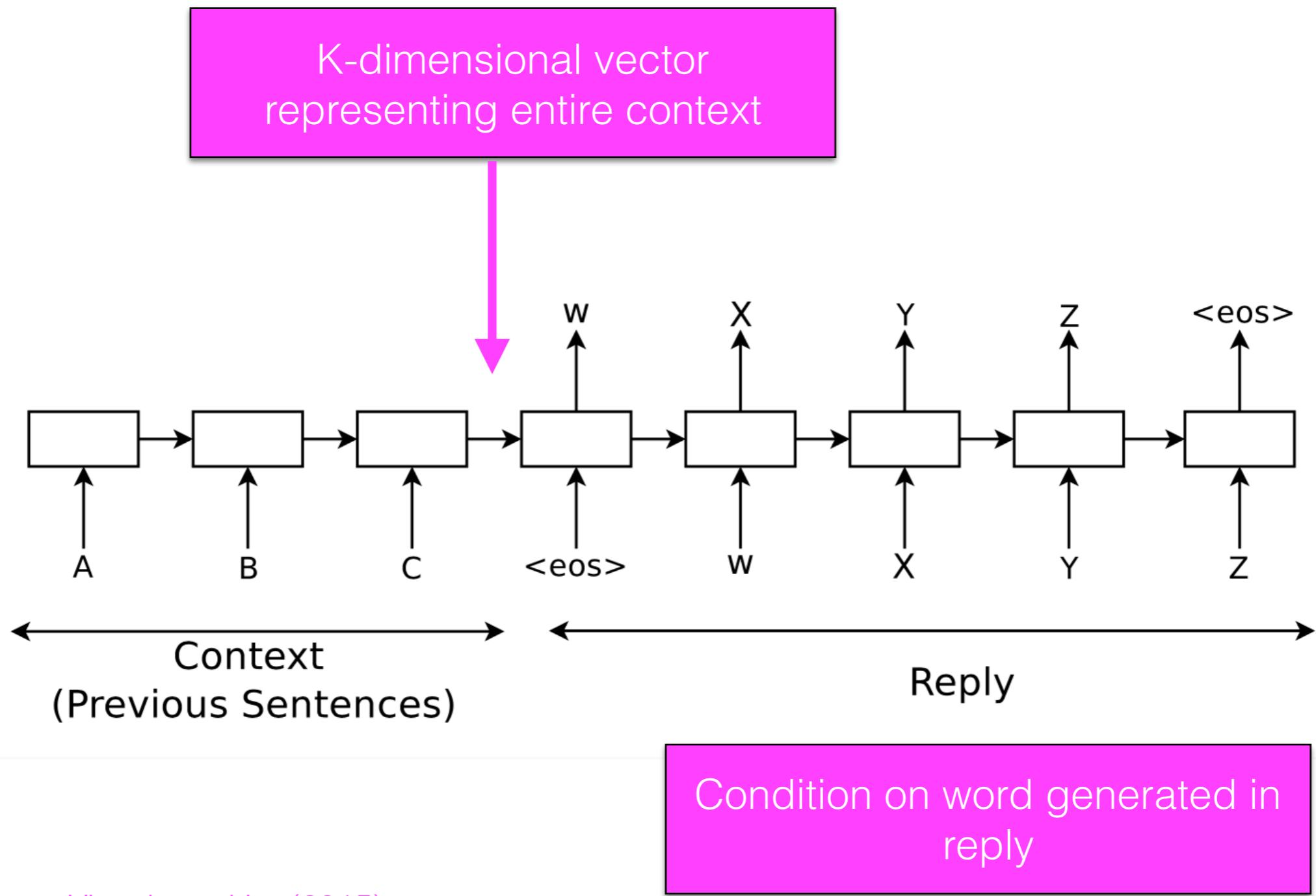
# Encoder-Decoder framework

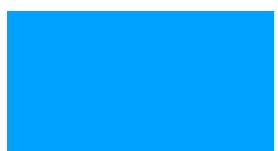


# Encoder-Decoder framework

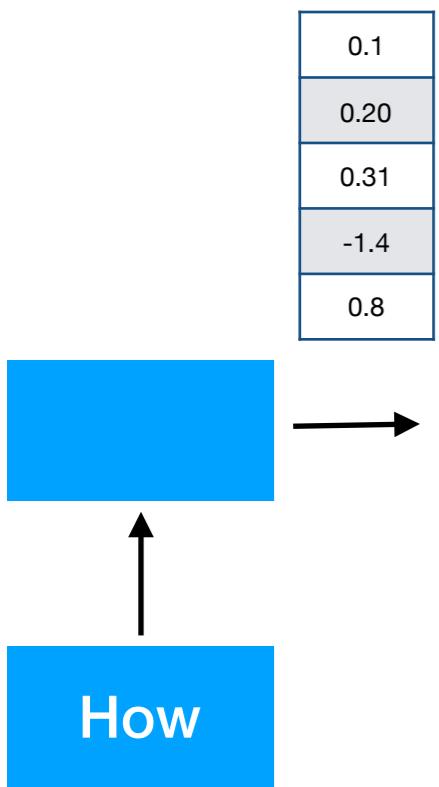


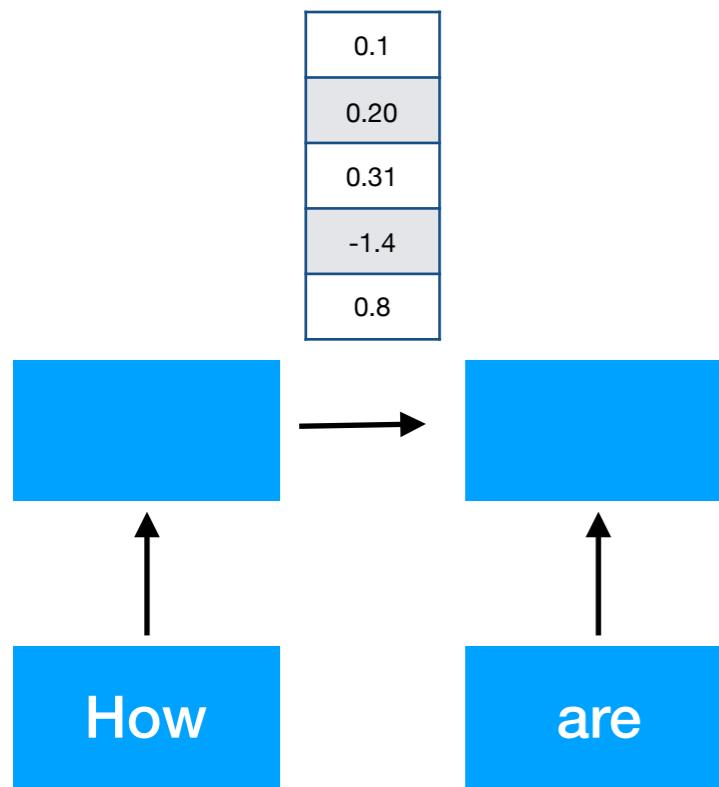
# Encoder-Decoder framework

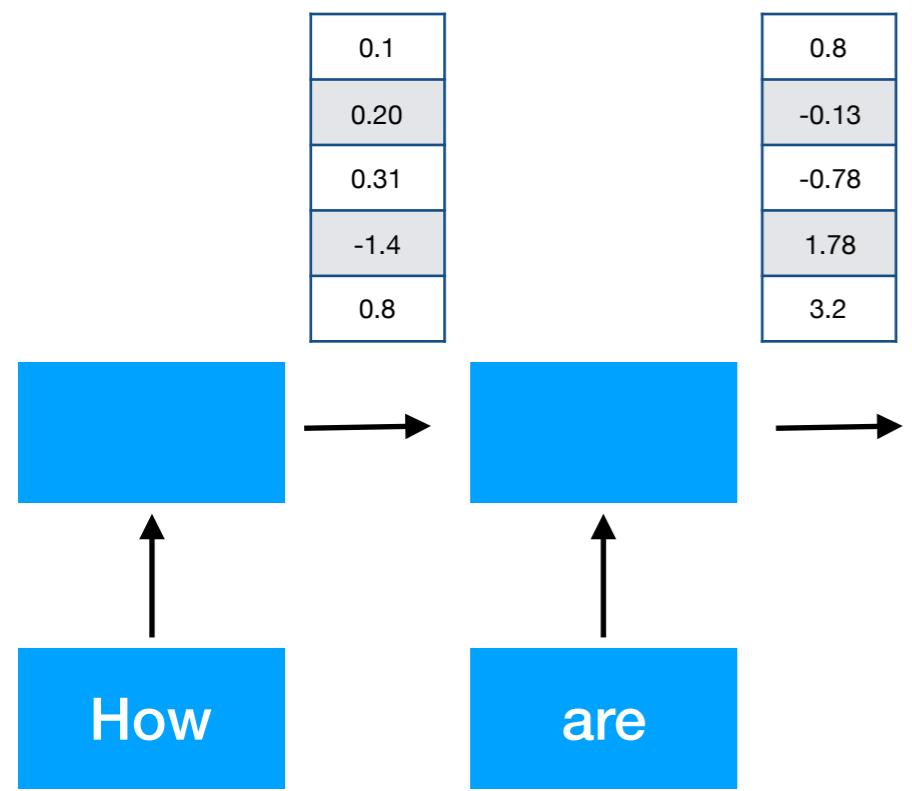


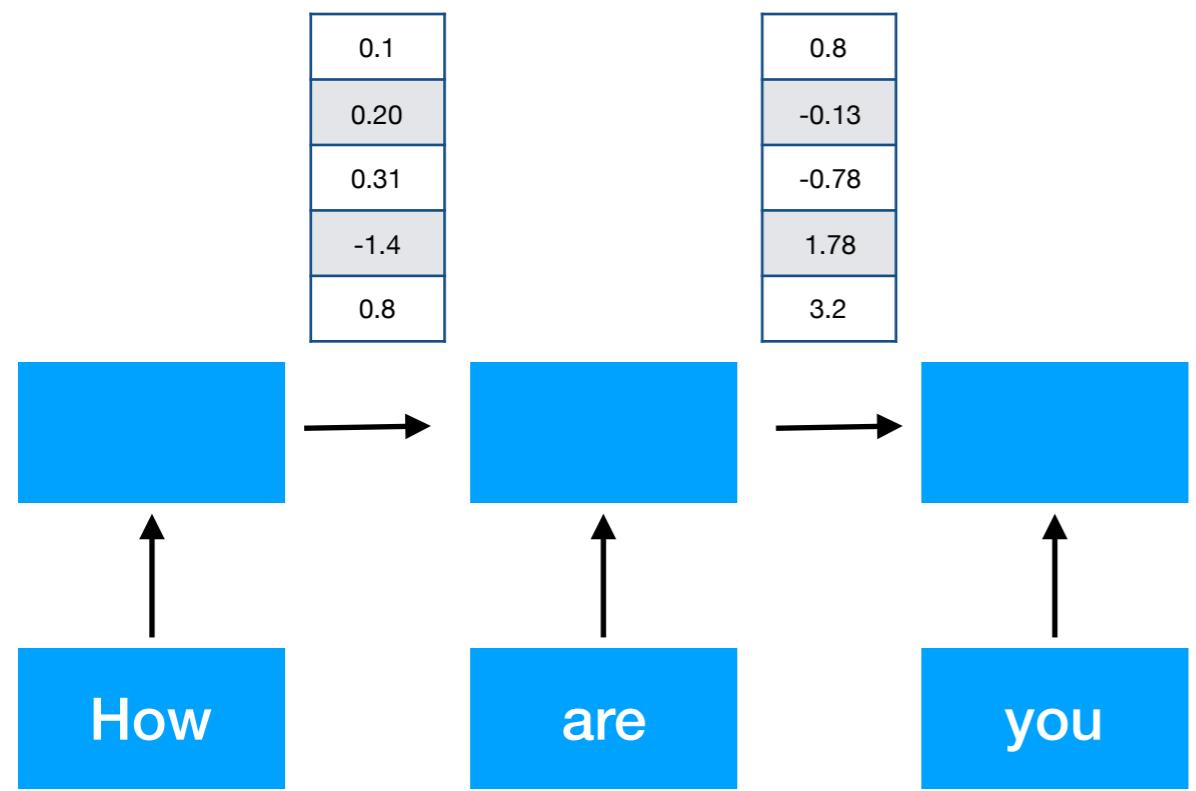


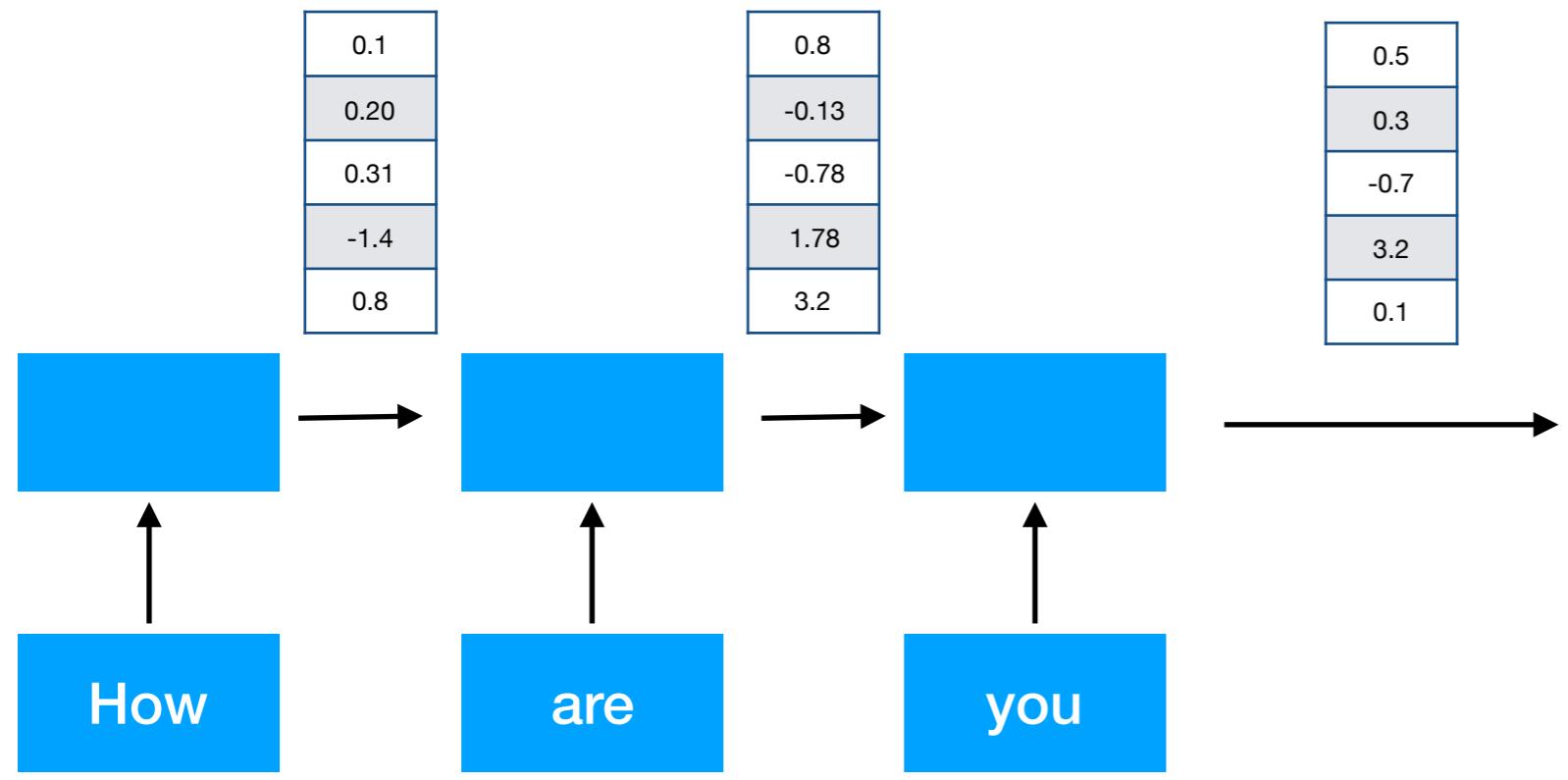
How

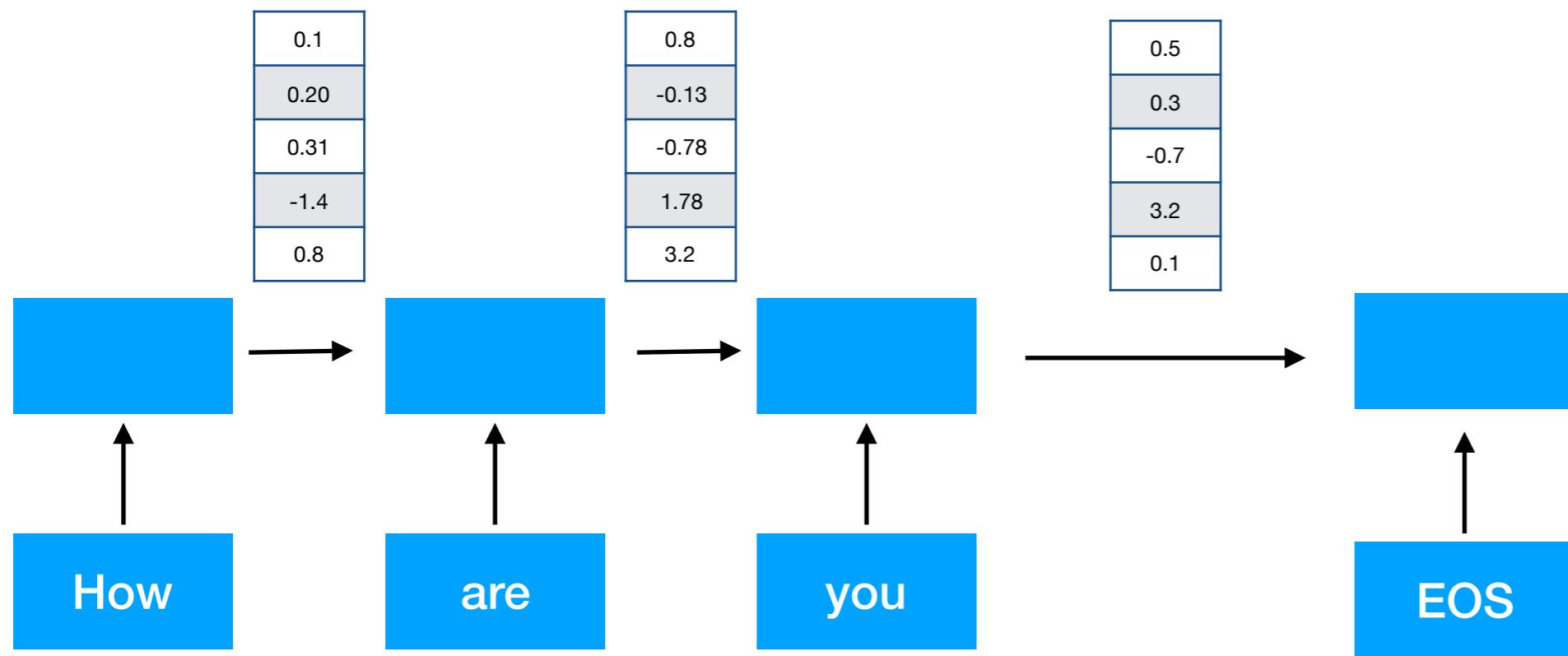


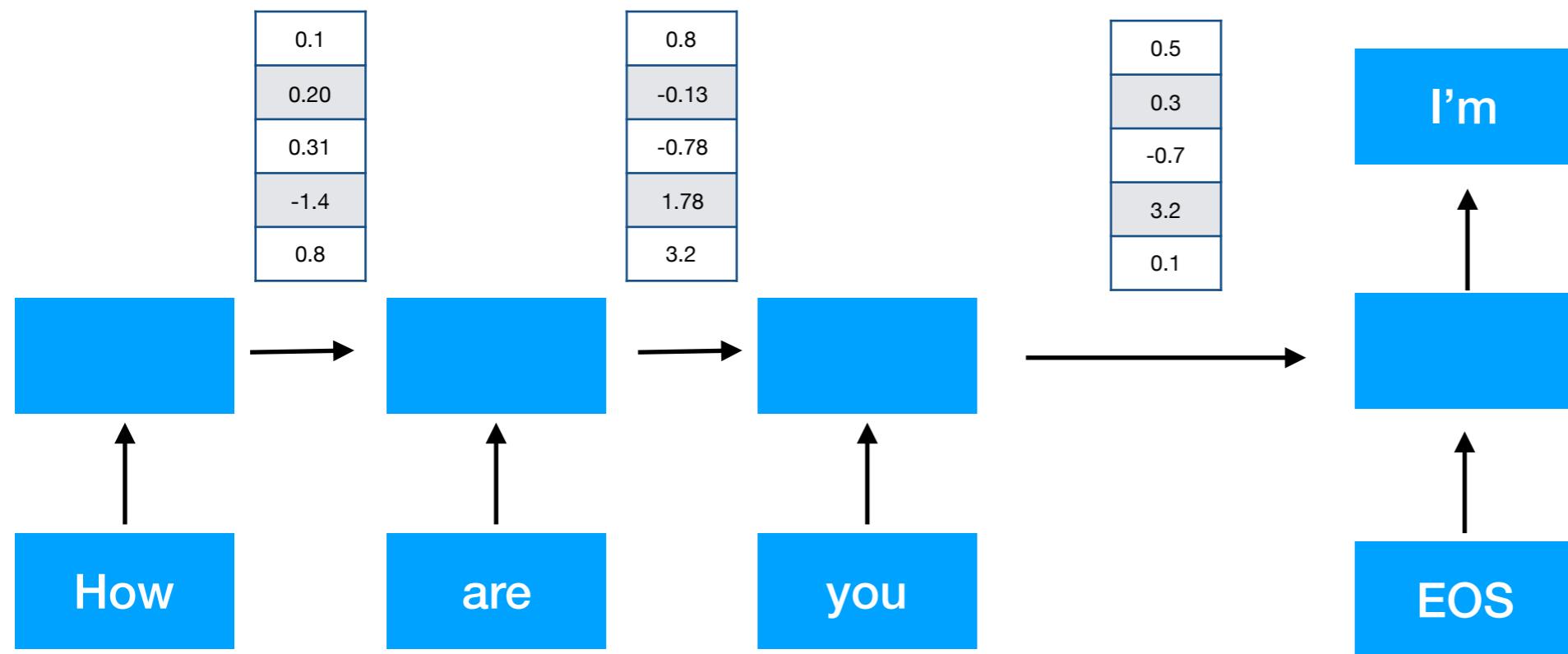


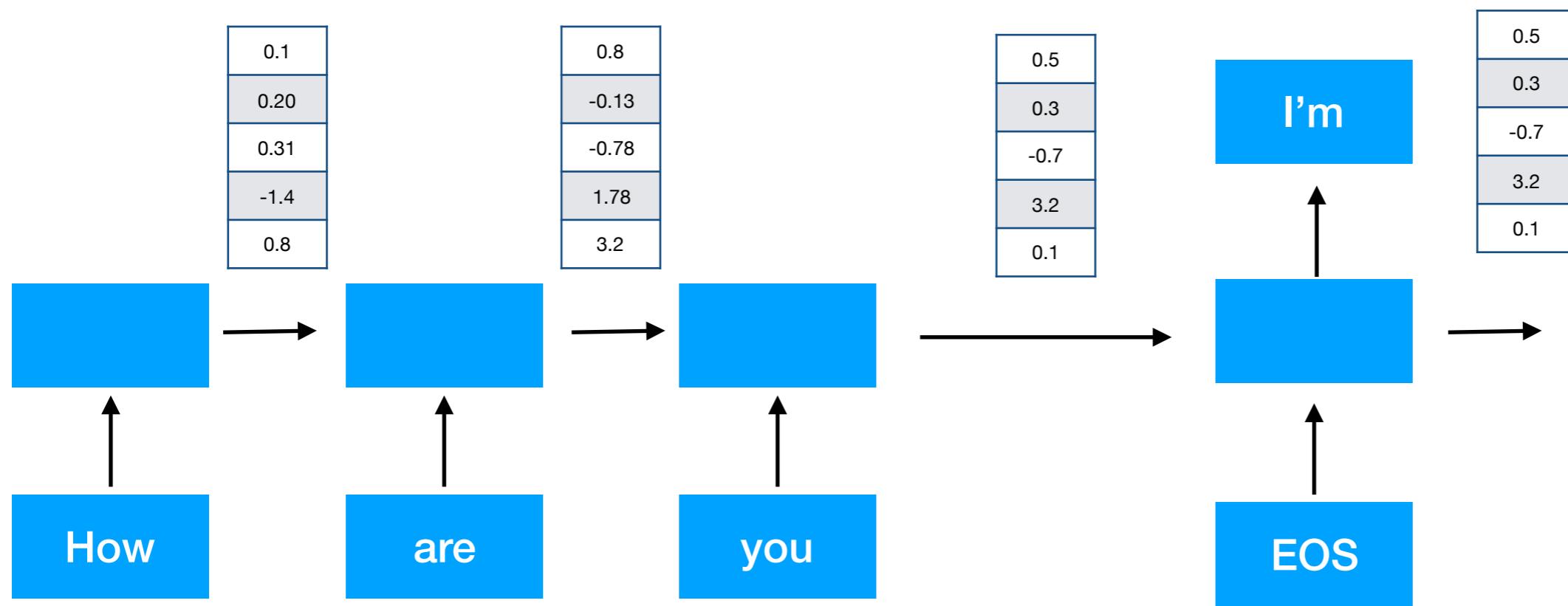


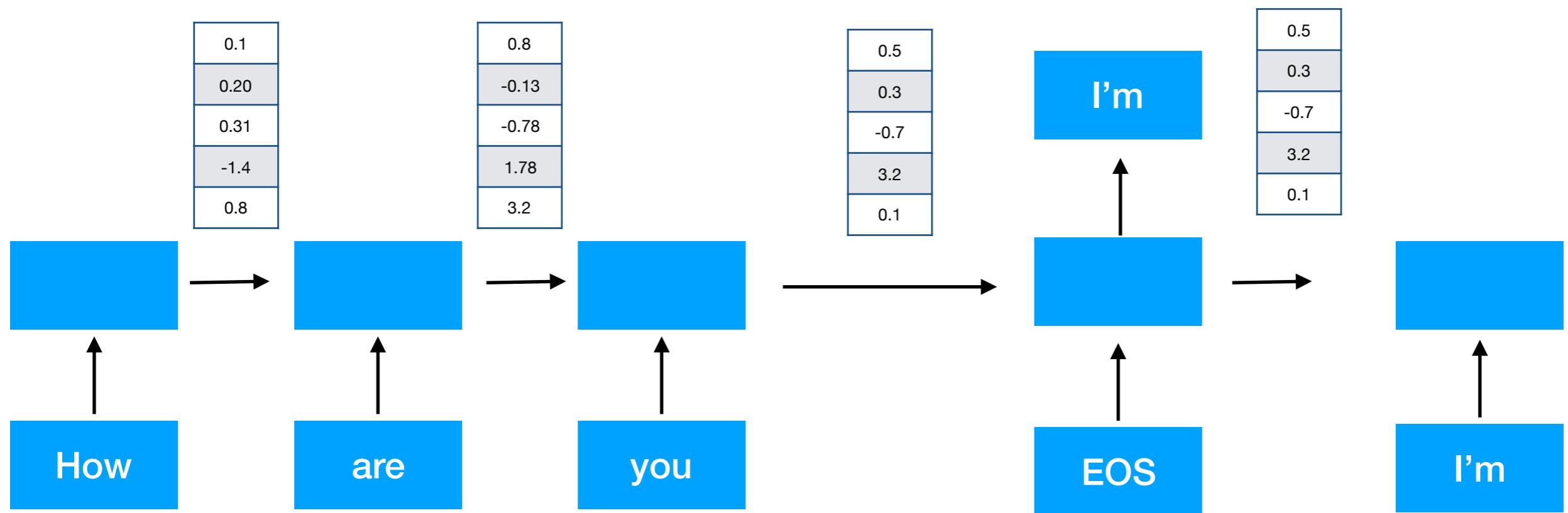


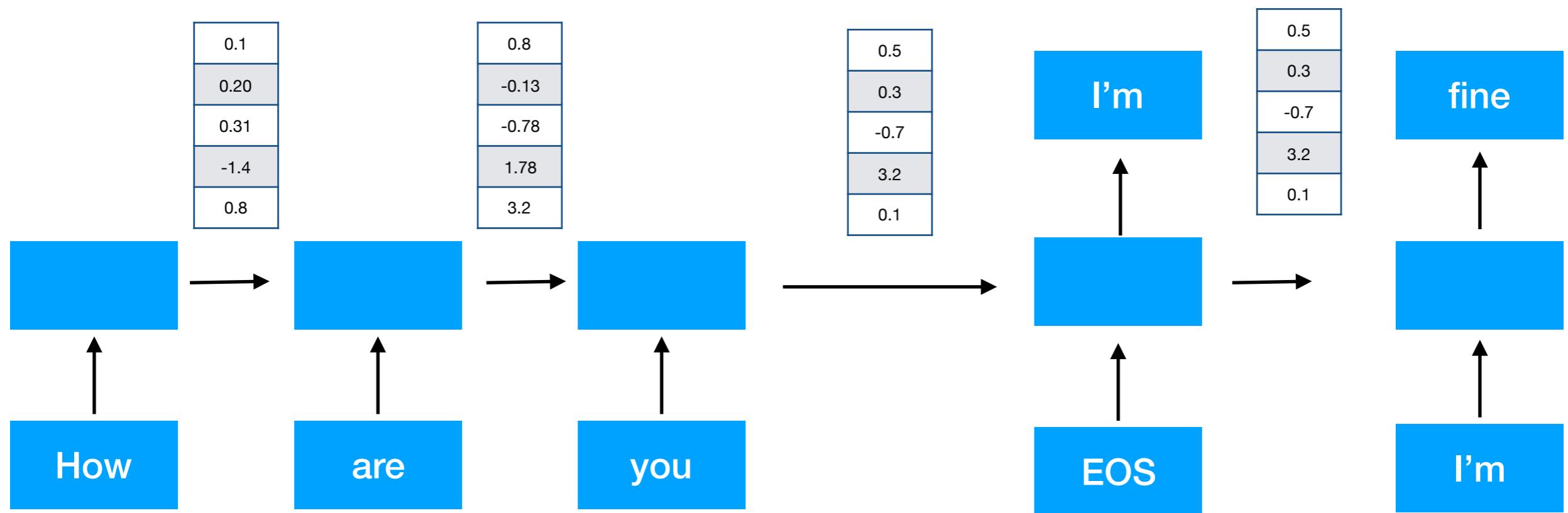


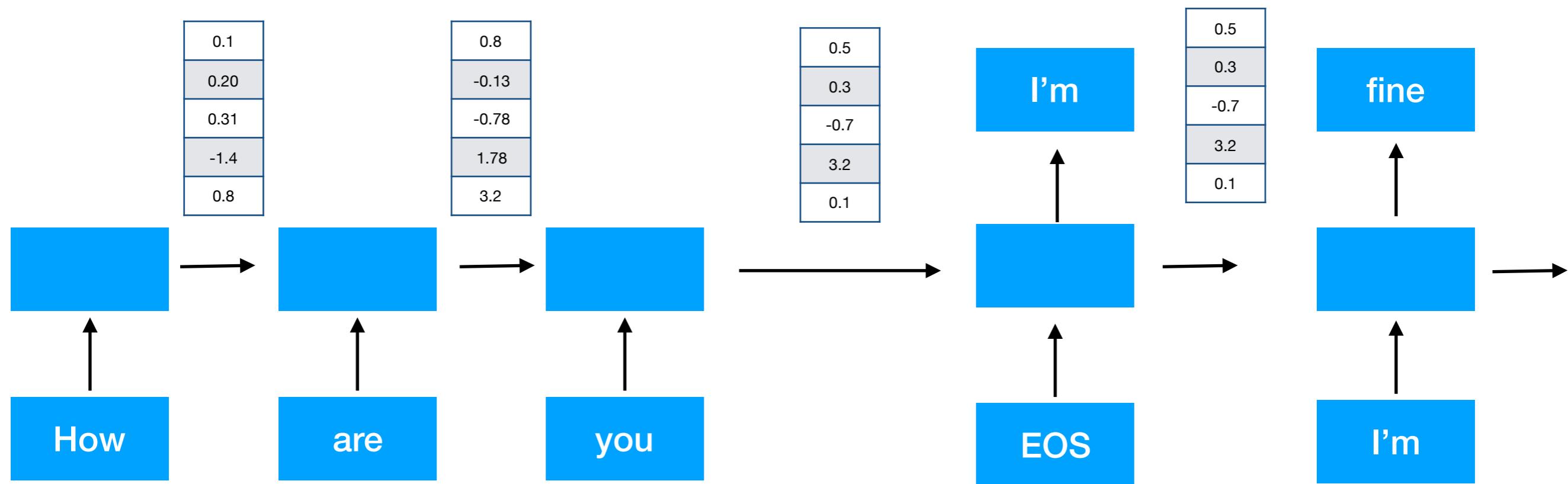






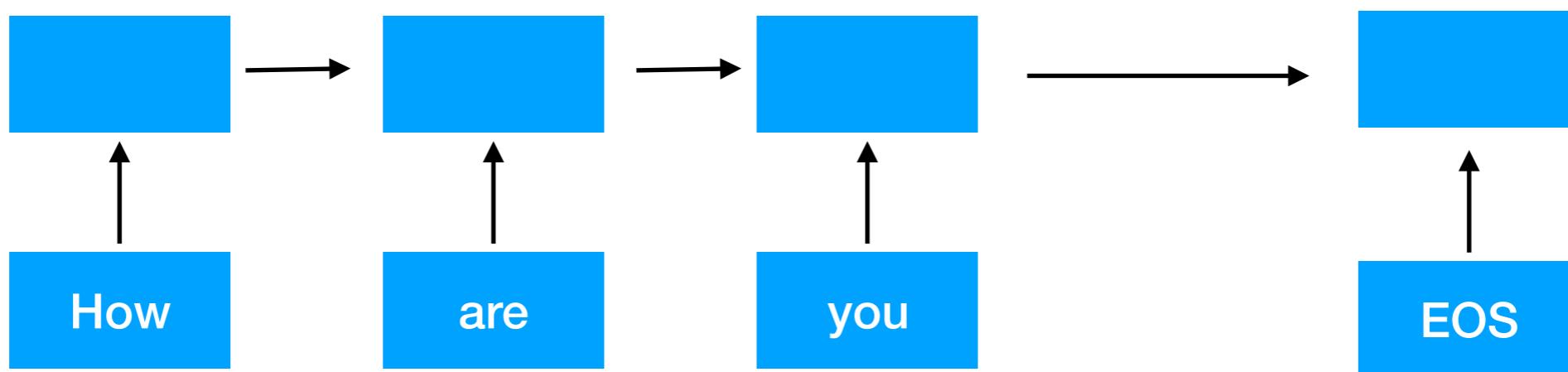






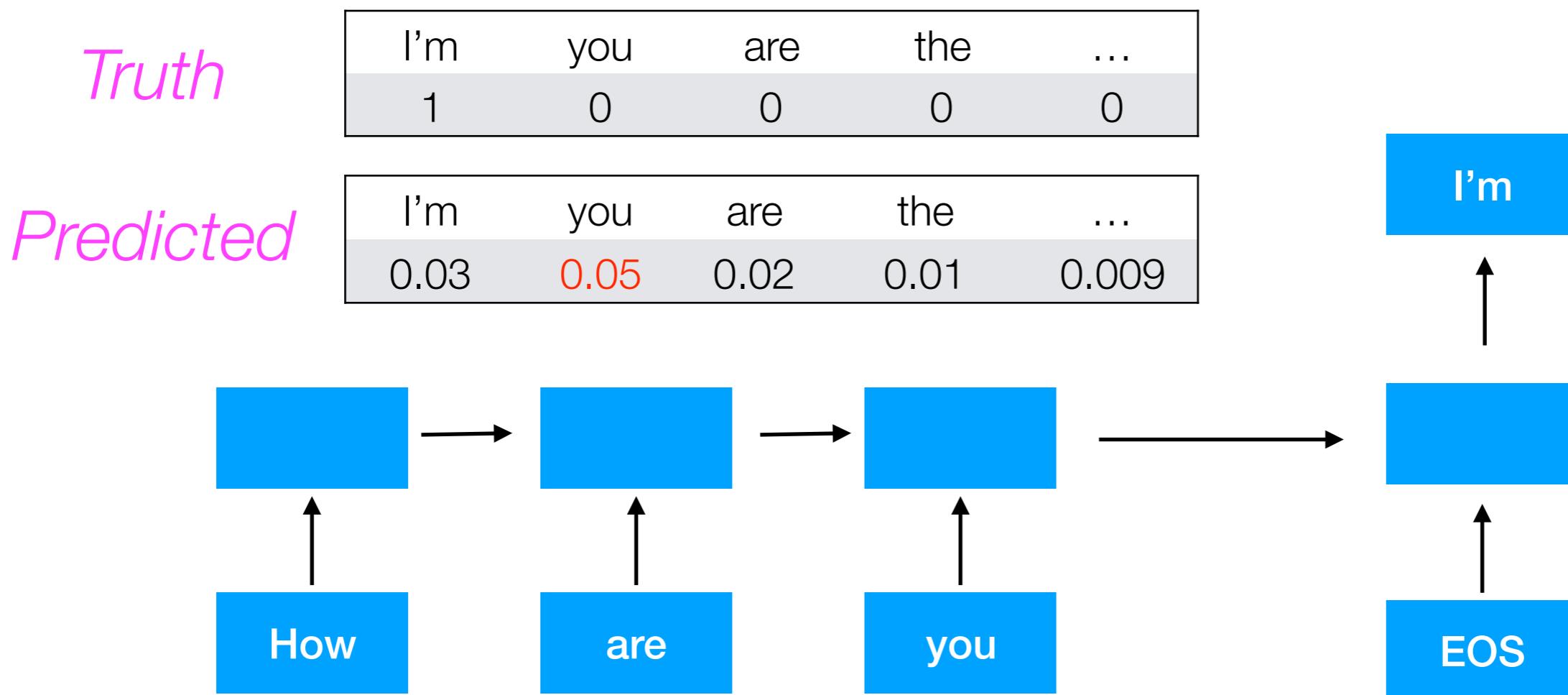
# Training

- As in other RNNs, we can train by minimizing the loss between what we predict at each time step and the truth.



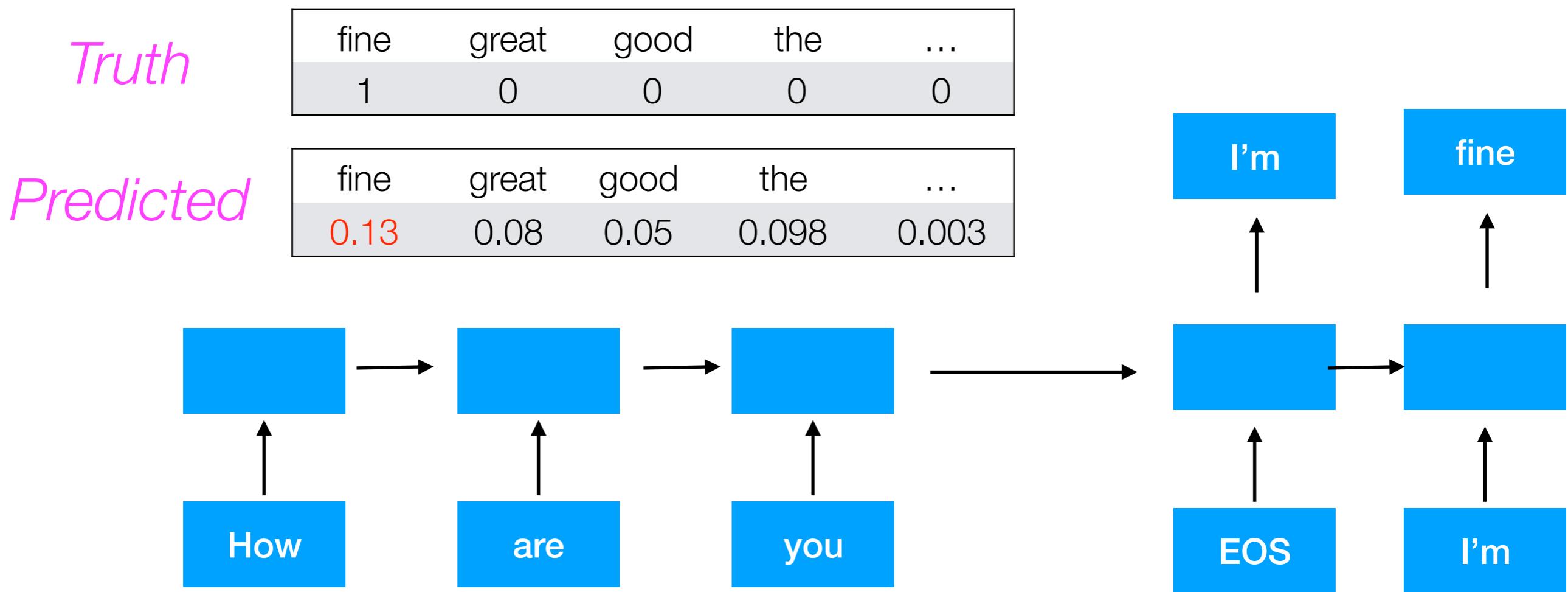
# Training

- As in other RNNs, we can train by minimizing the loss between what we predict at each time step and the truth.



# Training

- As in other RNNs, we can train by minimizing the loss between what we predict at each time step and the truth.



# Neural models

# Neural models

- Data: train on existing conversations

# Neural models

- Data: train on existing conversations
  - OpenSubtitles (movie conversations; 62M sentences/ 923M tokens). [Open domain](#). [Vinyals and Le 2015]

# Neural models

- Data: train on existing conversations
  - OpenSubtitles (movie conversations; 62M sentences/ 923M tokens). **Open domain.** [Vinyals and Le 2015]
  - Movie scripts (Friends/Big Bang Theory: dyadic interactions).

# Neural models

- Data: train on existing conversations
  - OpenSubtitles (movie conversations; 62M sentences/ 923M tokens). **Open domain.** [Vinyals and Le 2015]
  - Movie scripts (Friends/Big Bang Theory: dyadic interactions).
  - Twitter: minimum 3-turn conversations (context/ message/response); 24M sequences. [Li et al. 2016]

# Neural models

- Data: train on existing conversations
  - OpenSubtitles (movie conversations; 62M sentences/ 923M tokens). **Open domain.** [Vinyals and Le 2015]
  - Movie scripts (Friends/Big Bang Theory: dyadic interactions).
  - Twitter: minimum 3-turn conversations (context/ message/response); 24M sequences. [Li et al. 2016]
  - IT HelpDesk Troubleshooting data (30M tokens). **Narrow domain.** [Vinyals and Le 2015]

**Human:** *what is moral ?*

**Machine:** *what empowered humanity , what intellectual the essence is .*

**Human:** *what is immoral ?*

**Machine:** *the fact that you have a child .*

**Human:** *what is morality ?*

**Machine:** *what is altruism ?*

**Human:** *what is the definition of altruism ?*

**Machine:** *if you don 't believe in god , then you don 't know .*

**Human:** *ok ... so what is the definition of morality ?*

**Machine:** *well , the truth is , you 're not a believer in god almighty .*

**Human:** *tell me the definition of morality , i am quite upset now !*

**Machine:** *i 'm not ashamed of being a philosopher !*

**Human:** *tell me the definition of morality .*

**Machine:** *i don 't have ethics .*

--

# Evaluation

How do we evaluate conversational agents?

# Evaluation

- **Perplexity**: given a held-out dialogue response not used in training, how surprised are we by the words we see?

# Evaluation

- **BLEU** score: given a held-out dialogue response not used in training, how closely does a generated response match it (in terms of ngram overlap)?
- Not perfect because many responses are valid (unlike in machine translation where the space of possible translations for a fixed source is more constrained).

# Evaluation

- **Human judgment:**  
human judges to evaluation which of two conversational agents they prefer

**Human:** What is the deepest spot on the world ?  
**NCM:** The pacific ocean.  
**CB:** What ?

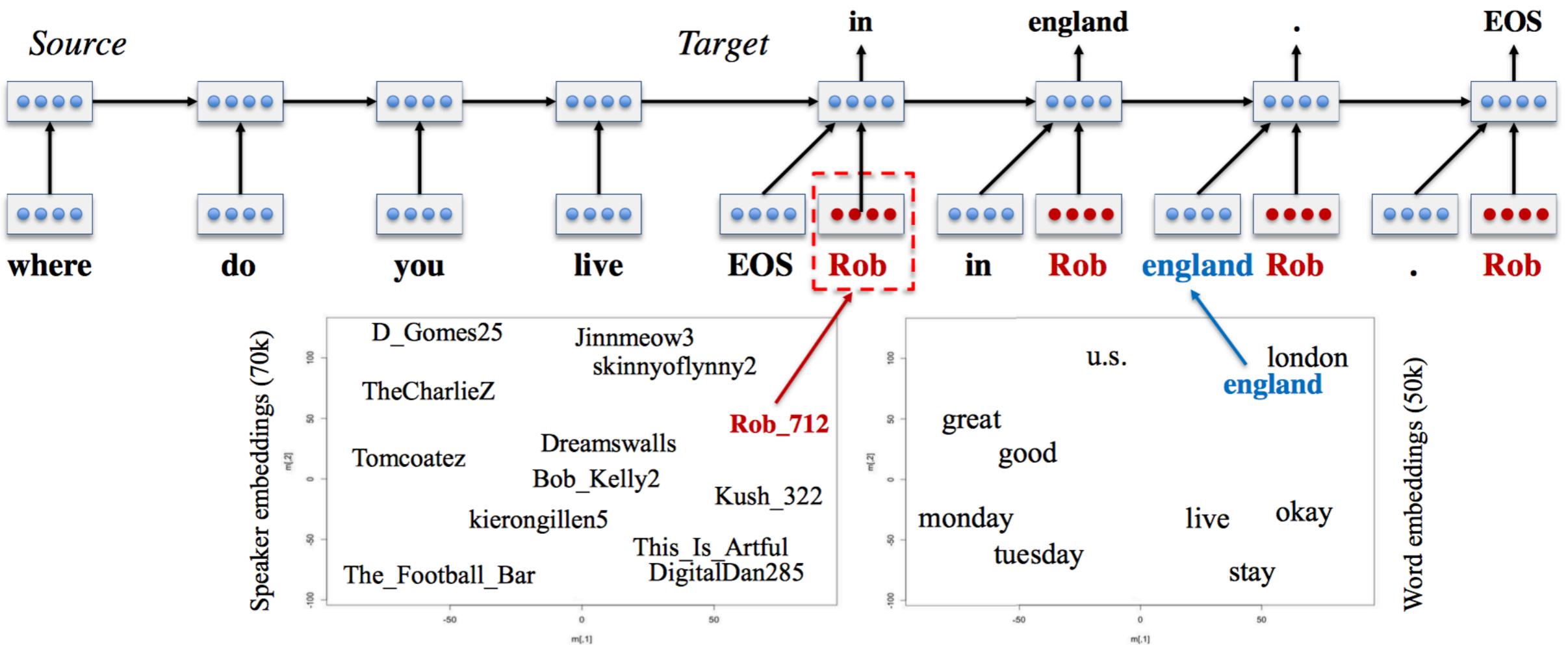
**Human:** Do you like Mexican food or Indian food ?  
**NCM:** I hate it.  
**CB:** I love Mexican food.

**Human:** Who are you crazy about ?  
**NCM:** You're the one who's crazy.  
**CB:** Who are you ?

# Personas

- We can model speaker-specific information (**latent** dialect, register, age, gender) to generate conversations under different personas
- Model this in a seq2seq model by conditioning on a k-dimensional representation of the user during generation.

# Personas



# Personas

- People also vary their dialogue according to the **addressee**.
- Model this in a seq2seq model by linearly combining user representation for speaker and addressee and conditioning response on that vector.

---

*message* Where is your hometown?

---

*baseline* I was born in Canada.

---

*user1* I'm from England.

*user2* I'm from Manchester.

*user3* I'm from Liverpool.

*user4* England. You?

*user5* My hometown.

*user6* I'm from Texas.

*user7* I'm from LA.

*user8* England.

*user9* I was born here.

*user10* I was born in the us.

---

---

*message* What company do you work for?

---

*baseline* I work for a company.

---

*user1* I don't have a job.

*user2* I'm a manager.

*user3* I'm working. #

*user4* I work for a company.

*user5* I'm working. #

*user6* Customer service.

*user7* I work in retail.

*user8* I work in a cafe.

*user9* I work at a gas station.

*user10* I don't know #

---

# Reinforcement learning

- Seq2seq models are trained to maximize  $P(\text{target} \mid \text{source})$
- This can prefer common stock phrases that are likely in any situation.

---

**Input:** What are you doing?

---

|       |                     |       |                     |
|-------|---------------------|-------|---------------------|
| -0.86 | I don't know.       | -1.09 | Get out of here.    |
| -1.03 | I don't know!       | -1.09 | I'm going home.     |
| -1.06 | Nothing.            | -1.09 | Oh my god!          |
| -1.09 | Get out of the way. | -1.10 | I'm talking to you. |

---

**Input:** what is your name?

---

|       |                    |       |                    |
|-------|--------------------|-------|--------------------|
| -0.91 | I don't know.      | ...   |                    |
| -0.92 | I don't know!      | -1.55 | My name is Robert. |
| -0.92 | I don't know, sir. | -1.58 | My name is John.   |
| -0.97 | Oh, my god!        | -1.59 | My name's John.    |

---

**Input:** How old are you?

---

|       |                |       |              |
|-------|----------------|-------|--------------|
| -0.79 | I don't know.  | ...   |              |
| -1.06 | I'm fine.      | -1.64 | Twenty-five. |
| -1.17 | I'm all right. | -1.66 | Five.        |
| -1.17 | I'm not sure.  | -1.71 | Eight.       |

---

| <b>Baseline mutual information model (Li et al. 2015)</b>                                                                                                                                                                                                                                            | <b>Proposed reinforcement learning model</b>                                                                                                                                                                                                                                                             |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>A: Where are you going? (1)</p> <p>B: I'm going to the restroom. (2)</p> <p>A: See you later. (3)</p> <p>B: See you later. (4)</p> <p>A: See you later. (5)</p> <p>B: See you later. (6)</p> <p>...</p> <p>...</p>                                                                                | <p>A: Where are you going? (1)</p> <p>B: I'm going to the police station. (2)</p> <p>A: I'll come with you. (3)</p> <p>B: No, no, no, no, you're not going anywhere. (4)</p> <p>A: Why? (5)</p> <p>B: I need you to stay here. (6)</p> <p>A: I don't know what you are talking about. (7)</p> <p>...</p> |
| <p>A: how old are you? (1)</p> <p>B: I'm 16. (2)</p> <p>A: 16? (3)</p> <p>B: I don't know what you are talking about. (4)</p> <p>A: You don't know what you are saying. (5)</p> <p>B: I don't know what you are talking about . (6)</p> <p>A: You don't know what you are saying. (7)</p> <p>...</p> | <p>A: How old are you? (1)</p> <p>B: I'm 16. Why are you asking? (2)</p> <p>A I thought you were 12. (3)</p> <p>B: What made you think so? (4)</p> <p>A: I don't know what you are talking about. (5)</p> <p>B: You don't know what you are saying. (6)</p> <p>...</p>                                   |

# Reinforcement learning

- A dyadic conversation takes place between two agents **p** and **q**.
- A conversation is a sequence of **actions** taken by the agents according to a **policy** defined by a seq2seq model.
- Parameters optimized to maximize the expected future reward (**over the entire conversation**)

Li et al. (2016), "Deep Reinforcement Learning for Dialogue Generation" (EMNLP)



more machine learning

# Quick Detour: Reinforcement Learning

# Markov Decision Processes

# Markov Decision Processes

Defined by:

# Markov Decision Processes

Defined by:

- A set of states  $s \in S$

# Markov Decision Processes

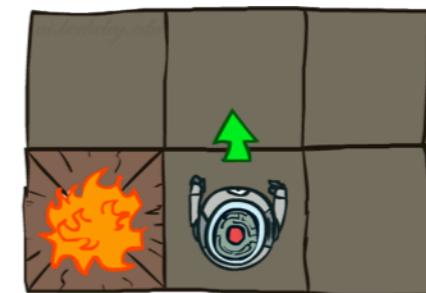
Defined by:

- A set of states  $s \in S$
- A set of actions  $a \in A$

# Markov Decision Processes

Defined by:

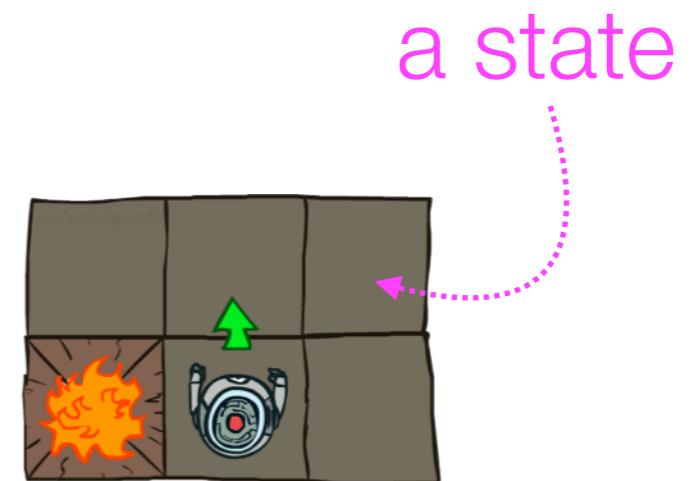
- A set of states  $s \in S$
- A set of actions  $a \in A$



# Markov Decision Processes

Defined by:

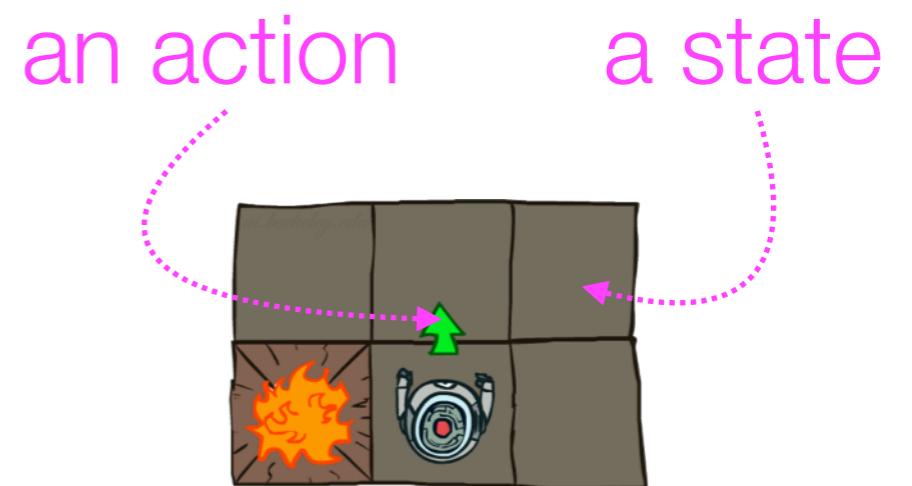
- A set of states  $s \in S$
- A set of actions  $a \in A$



# Markov Decision Processes

Defined by:

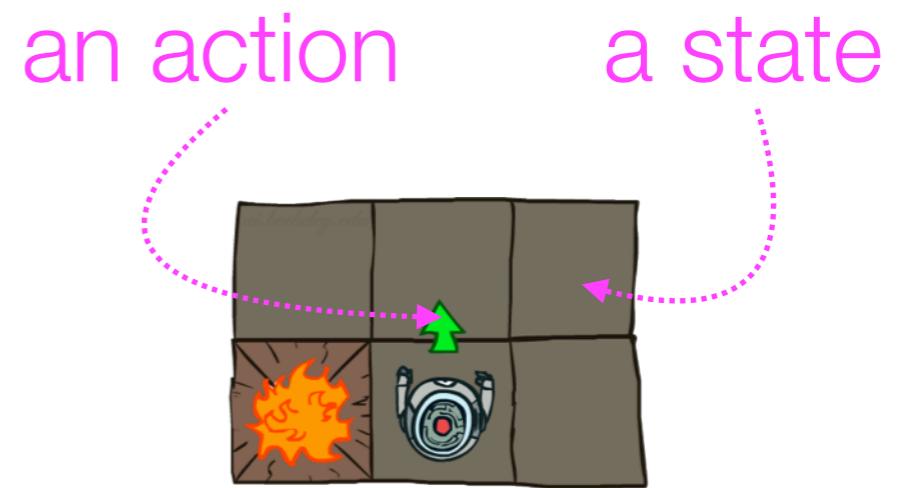
- A set of states  $s \in S$
- A set of actions  $a \in A$



# Markov Decision Processes

Defined by:

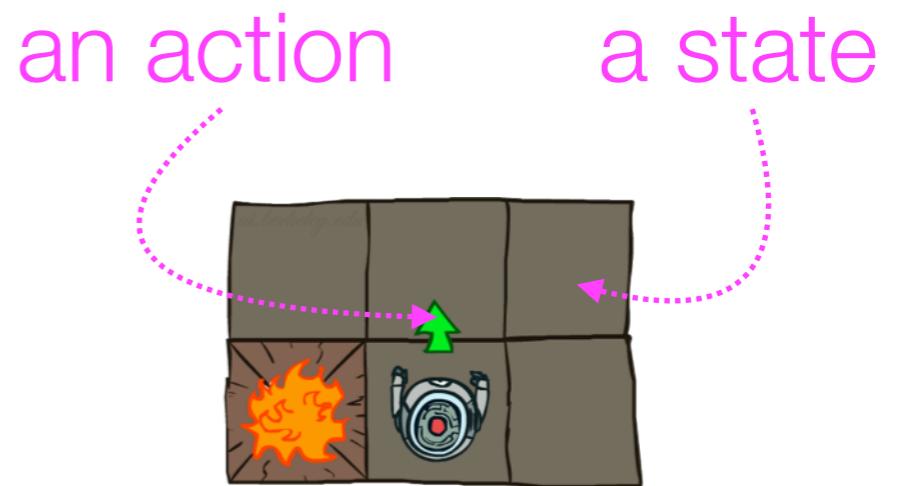
- A set of states  $s \in S$
- A set of actions  $a \in A$
- A transition function  $T(s, a, s')$ 
  - Probability that  $a$  from  $s$  leads to  $s'$ ,  
i.e.,  $P(s'|s, a)$



# Markov Decision Processes

Defined by:

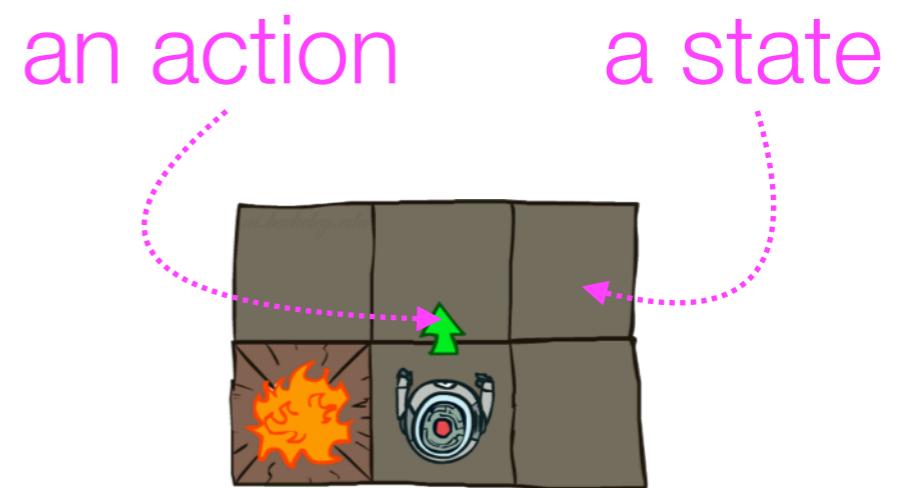
- A set of states  $s \in S$
- A set of actions  $a \in A$
- A transition function  $T(s, a, s')$ 
  - Probability that  $a$  from  $s$  leads to  $s'$ , i.e.,  $P(s'|s, a)$
- A reward function  $R(s, a, s')$ 
  - Sometimes just  $R(s)$  or  $R(s')$



# Markov Decision Processes

Defined by:

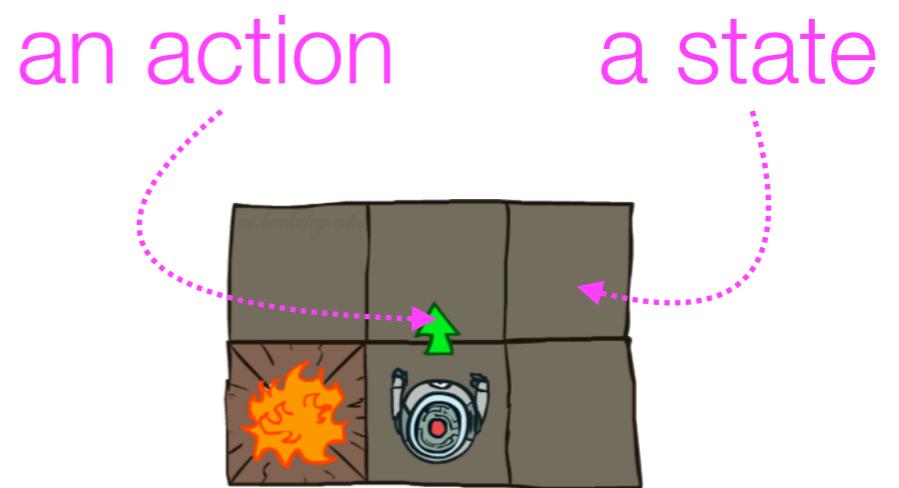
- A set of states  $s \in S$
- A set of actions  $a \in A$
- A transition function  $T(s, a, s')$ 
  - Probability that  $a$  from  $s$  leads to  $s'$ ,  
i.e.,  $P(s'|s, a)$
- A reward function  $R(s, a, s')$ 
  - Sometimes just  $R(s)$  or  $R(s')$
- A start state



# Markov Decision Processes

Defined by:

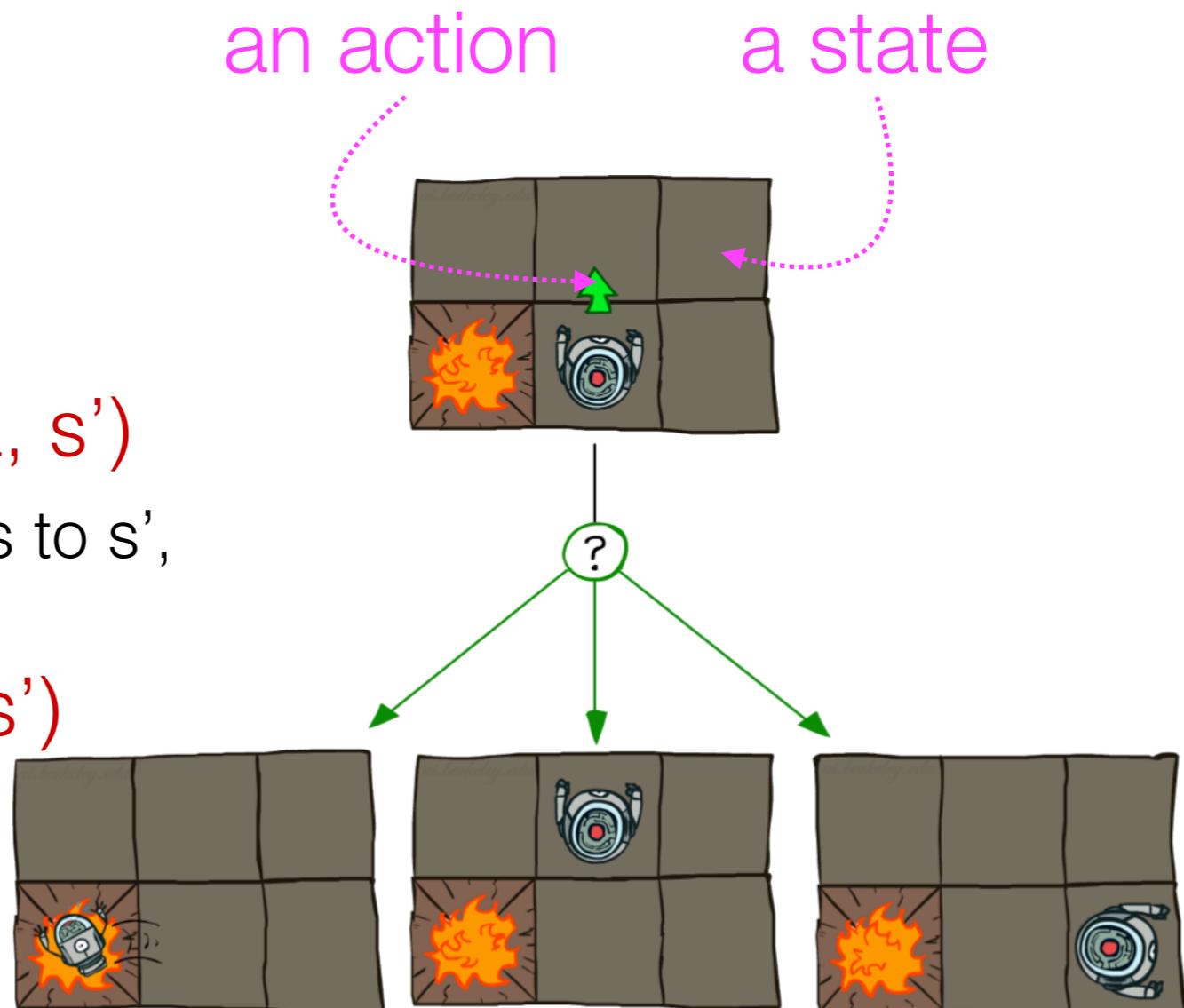
- A set of states  $s \in S$
- A set of actions  $a \in A$
- A transition function  $T(s, a, s')$ 
  - Probability that  $a$  from  $s$  leads to  $s'$ ,  
i.e.,  $P(s'|s, a)$
- A reward function  $R(s, a, s')$ 
  - Sometimes just  $R(s)$  or  $R(s')$
- A start state
- Maybe a terminal state



# Markov Decision Processes

Defined by:

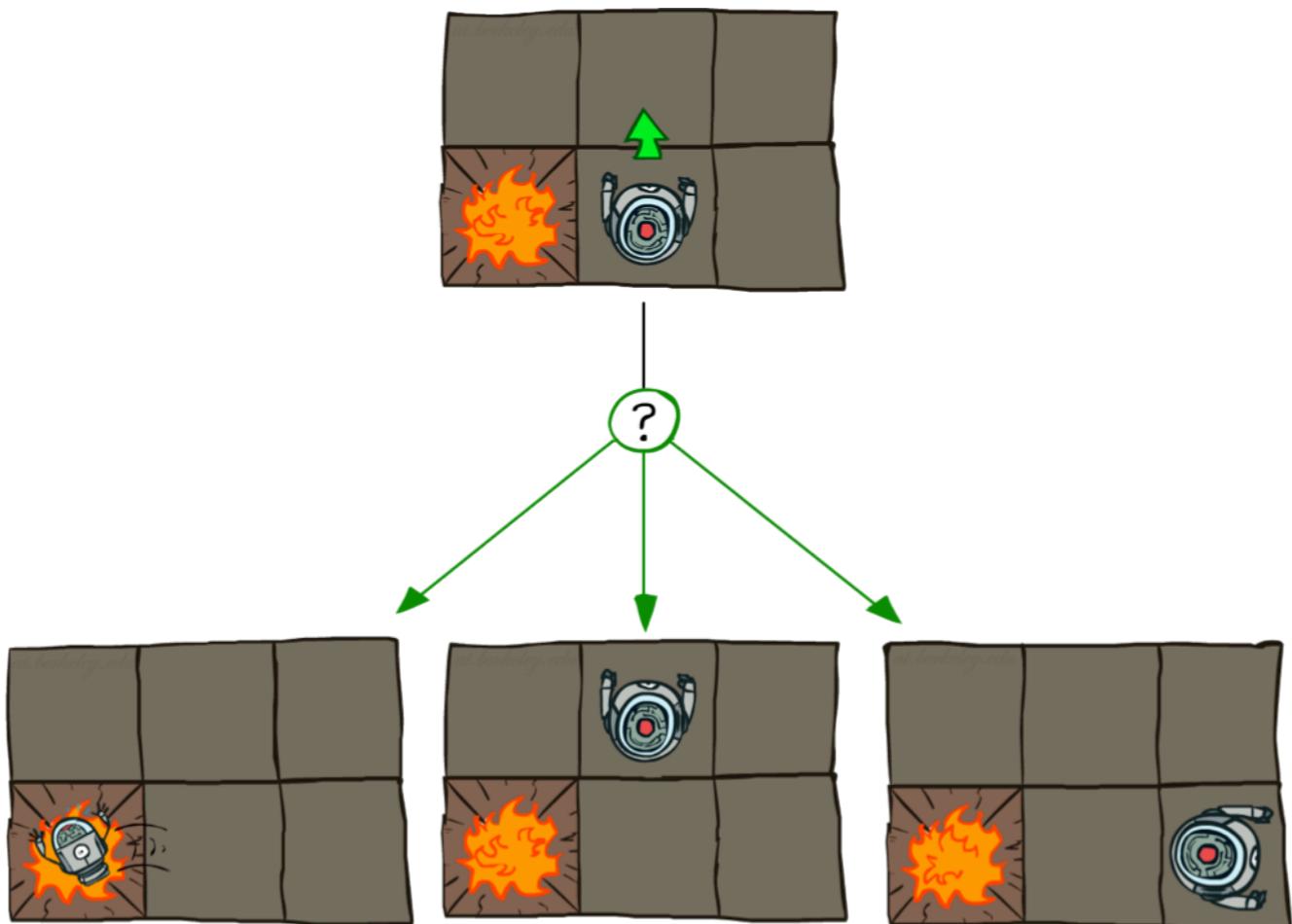
- A set of states  $s \in S$
- A set of actions  $a \in A$
- A transition function  $T(s, a, s')$ 
  - Probability that  $a$  from  $s$  leads to  $s'$ , i.e.,  $P(s'|s, a)$
- A reward function  $R(s, a, s')$ 
  - Sometimes just  $R(s)$  or  $R(s')$
- A start state
- Maybe a terminal state



# Reinforcement Learning

Defined by:

- A set of states  $s \in S$
- A set of actions  $a \in A$
- A start state
- Maybe a terminal state

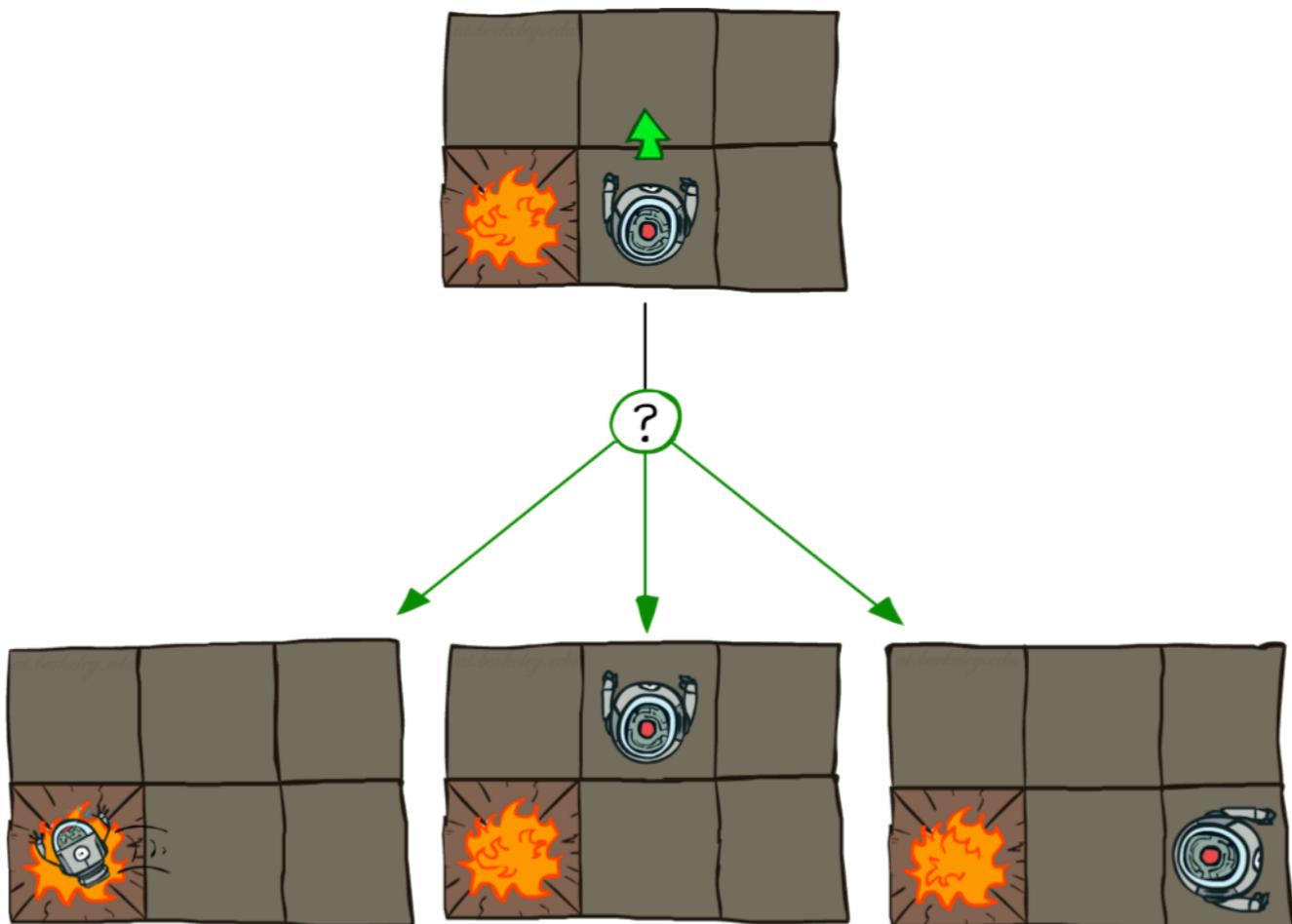


# Reinforcement Learning

Defined by:

- A set of states  $s \in S$
- A set of actions  $a \in A$
- A start state
- Maybe a terminal state

No longer know  
transition function  $T(s, a, s')$   
reward function  $R(s)$



# Reinforcement Learning

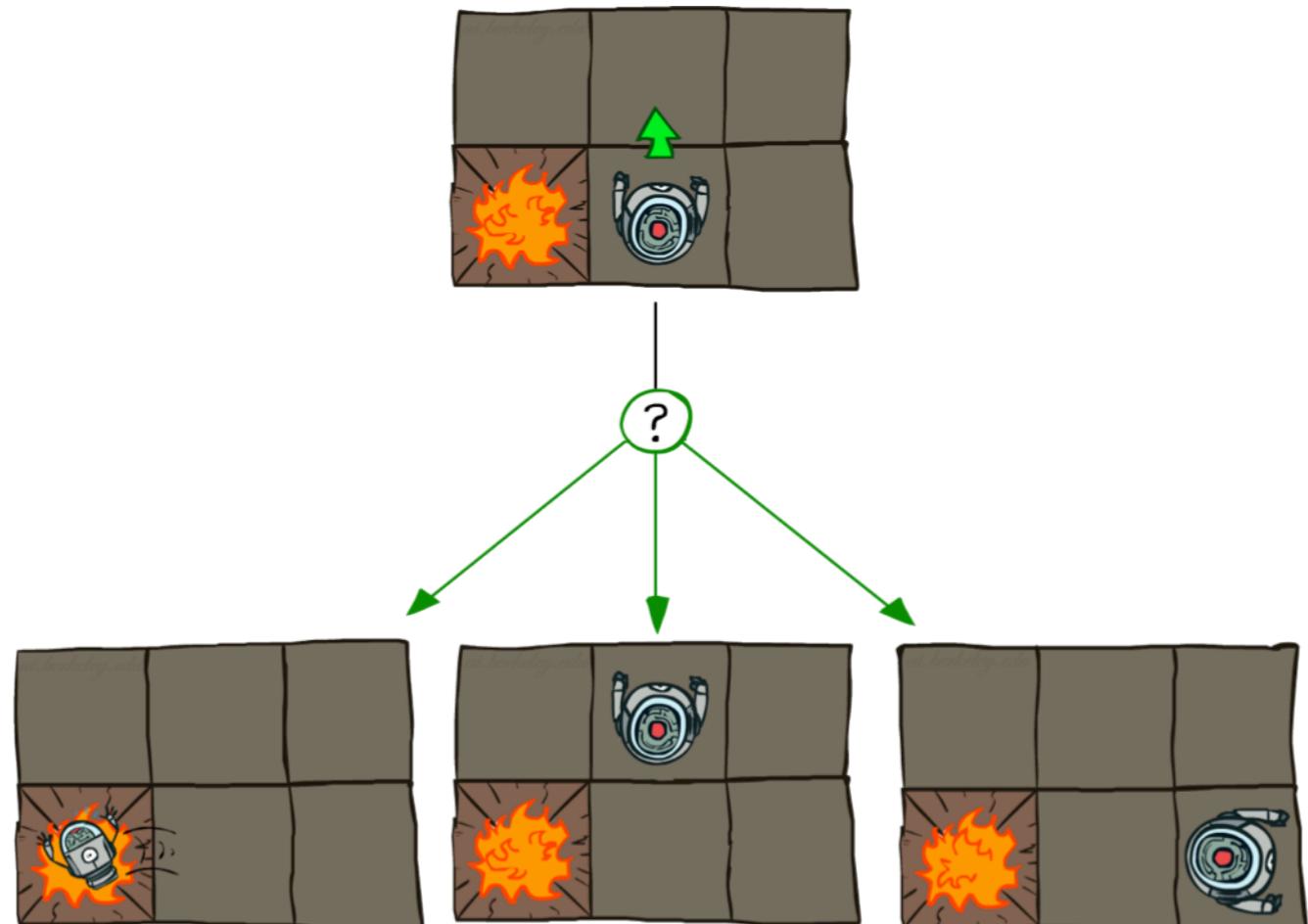
Defined by:

- A set of states  $s \in S$
- A set of actions  $a \in A$
- A start state
- Maybe a terminal state

No longer know  
transition function  $T(s, a, s')$   
reward function  $R(s)$

But, we do observe two things:

- End state  $s'$  after taking action  $a$  from state  $s$
- Reward  $R(s')$  when we enter  $s'$



# Reinforcement Learning

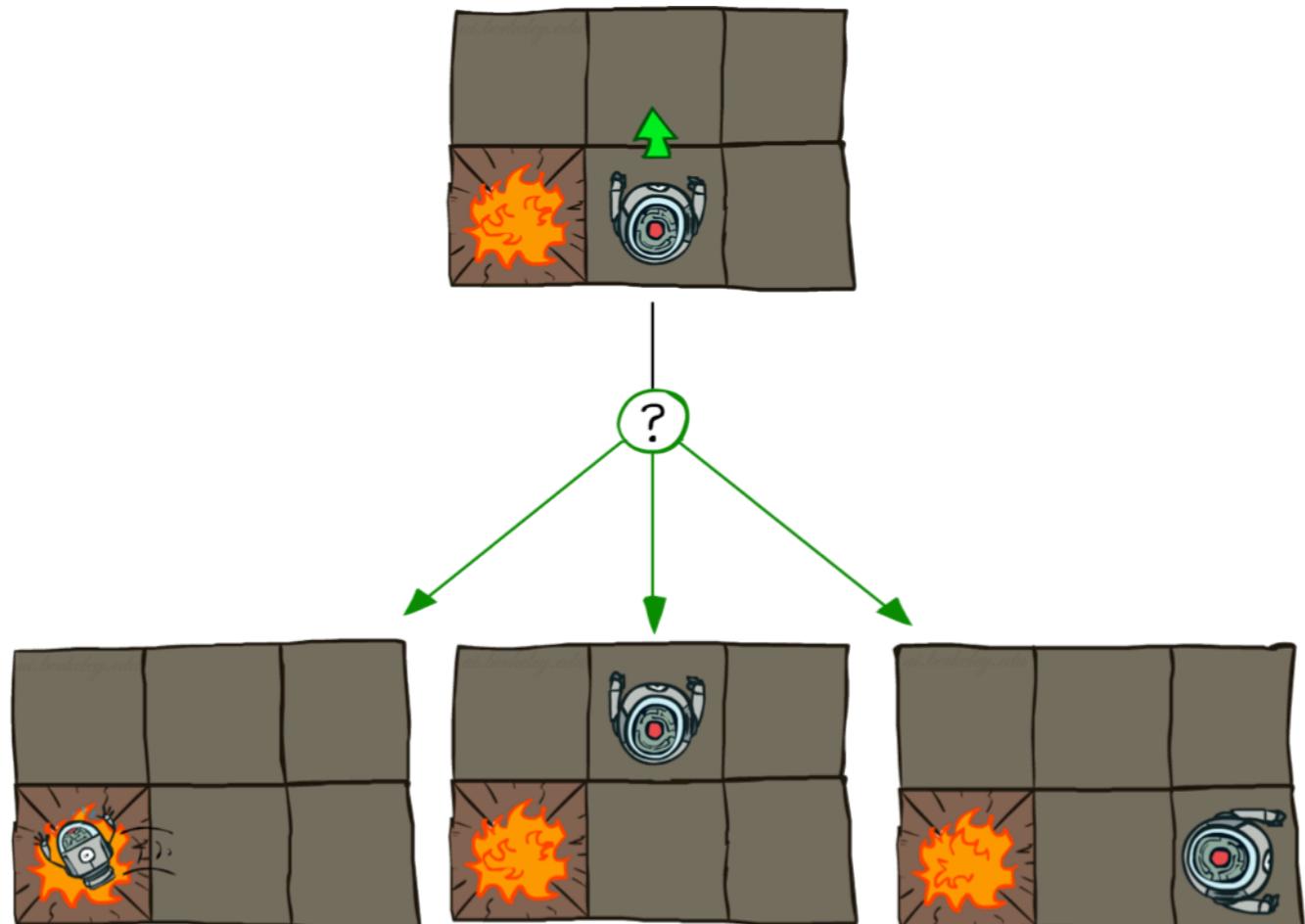
Defined by:

- A set of states  $s \in S$
- A set of actions  $a \in A$
- A start state
- Maybe a terminal state

No longer know  
transition function  $T(s, a, s')$   
reward function  $R(s)$

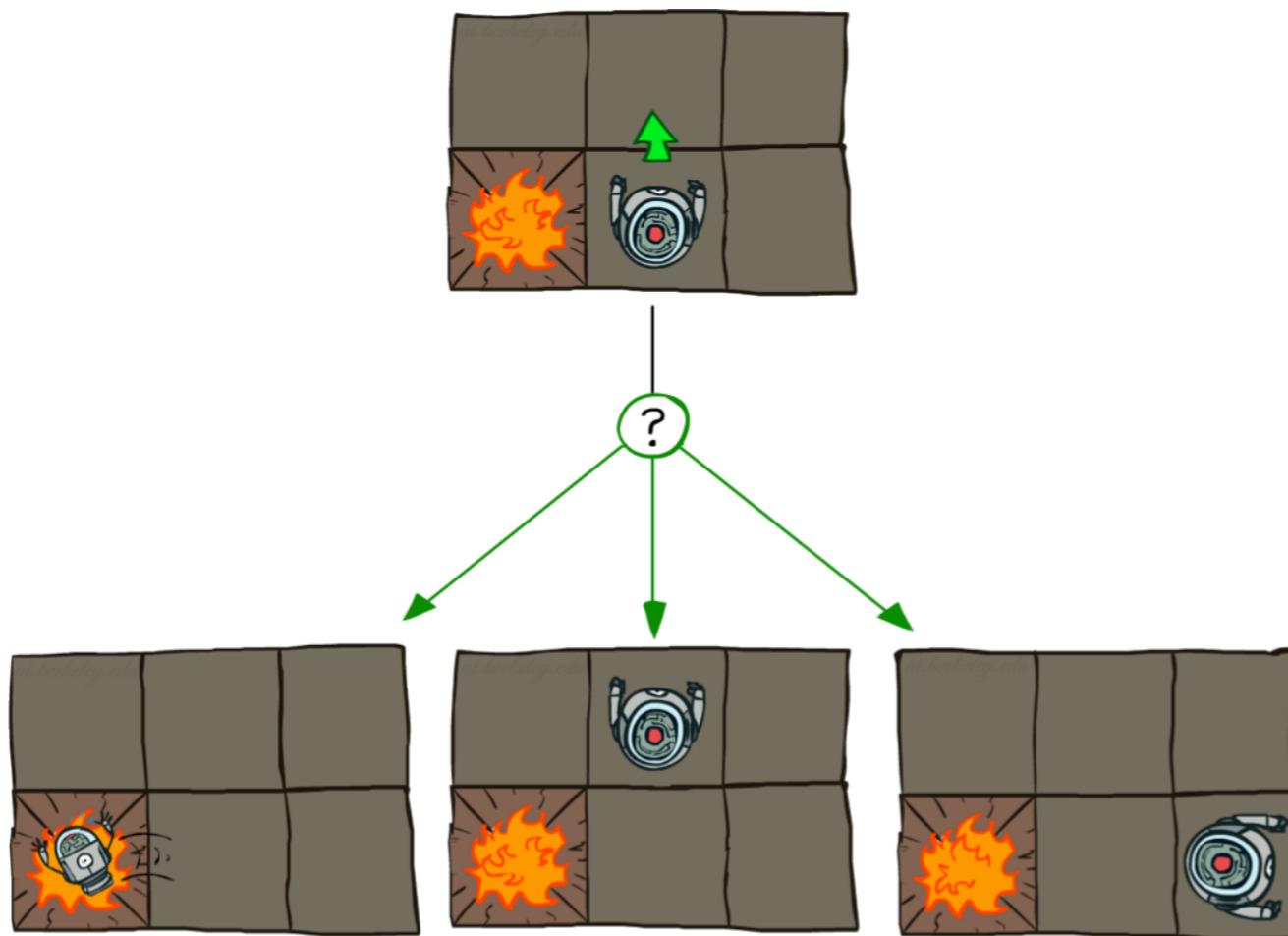
But, we do observe two things:

- End state  $s'$  after taking action  $a$  from state  $s$
- Reward  $R(s')$  when we enter  $s'$

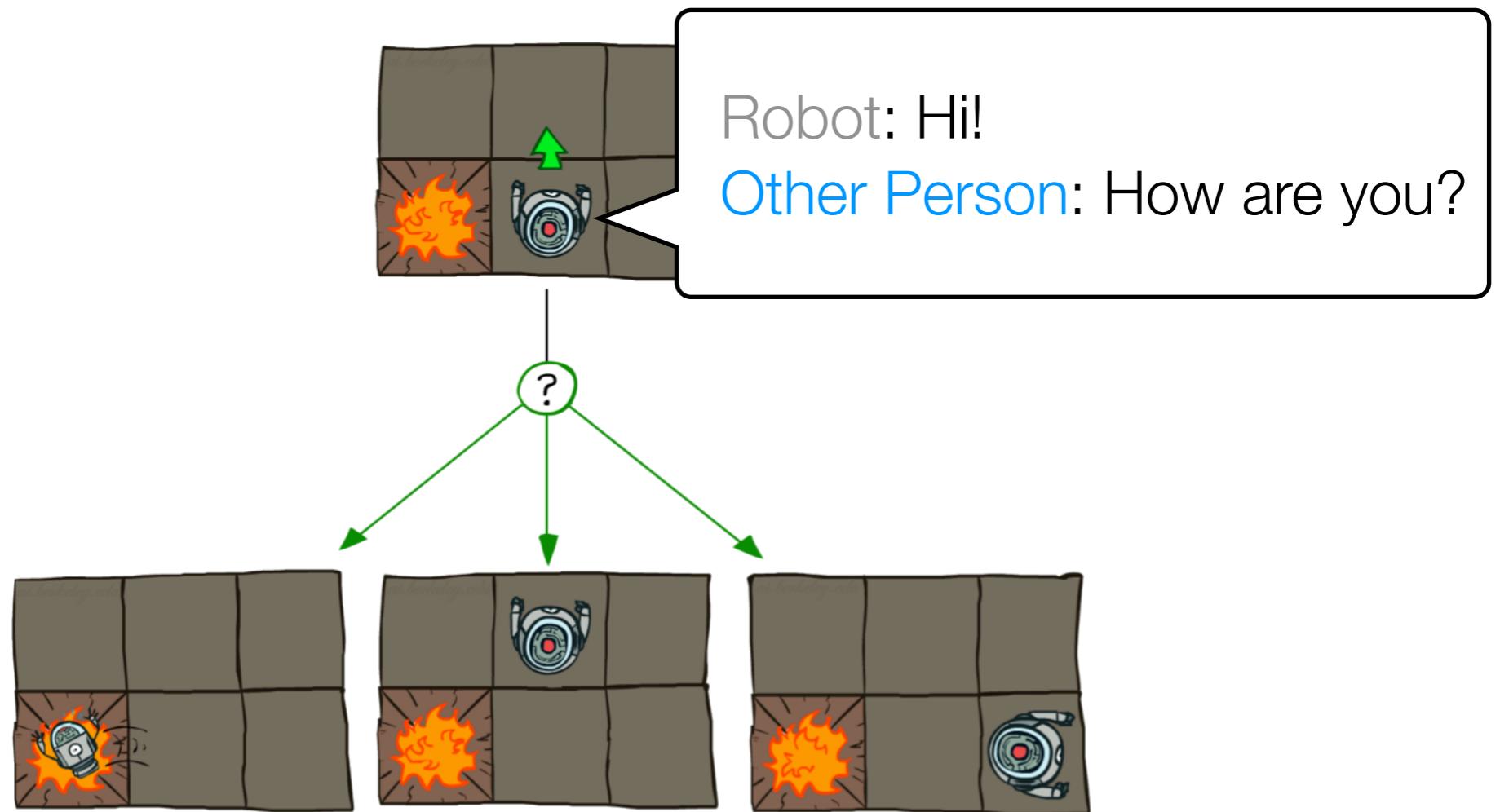


How would you use  
these to *learn* a decision  
policy for any state?

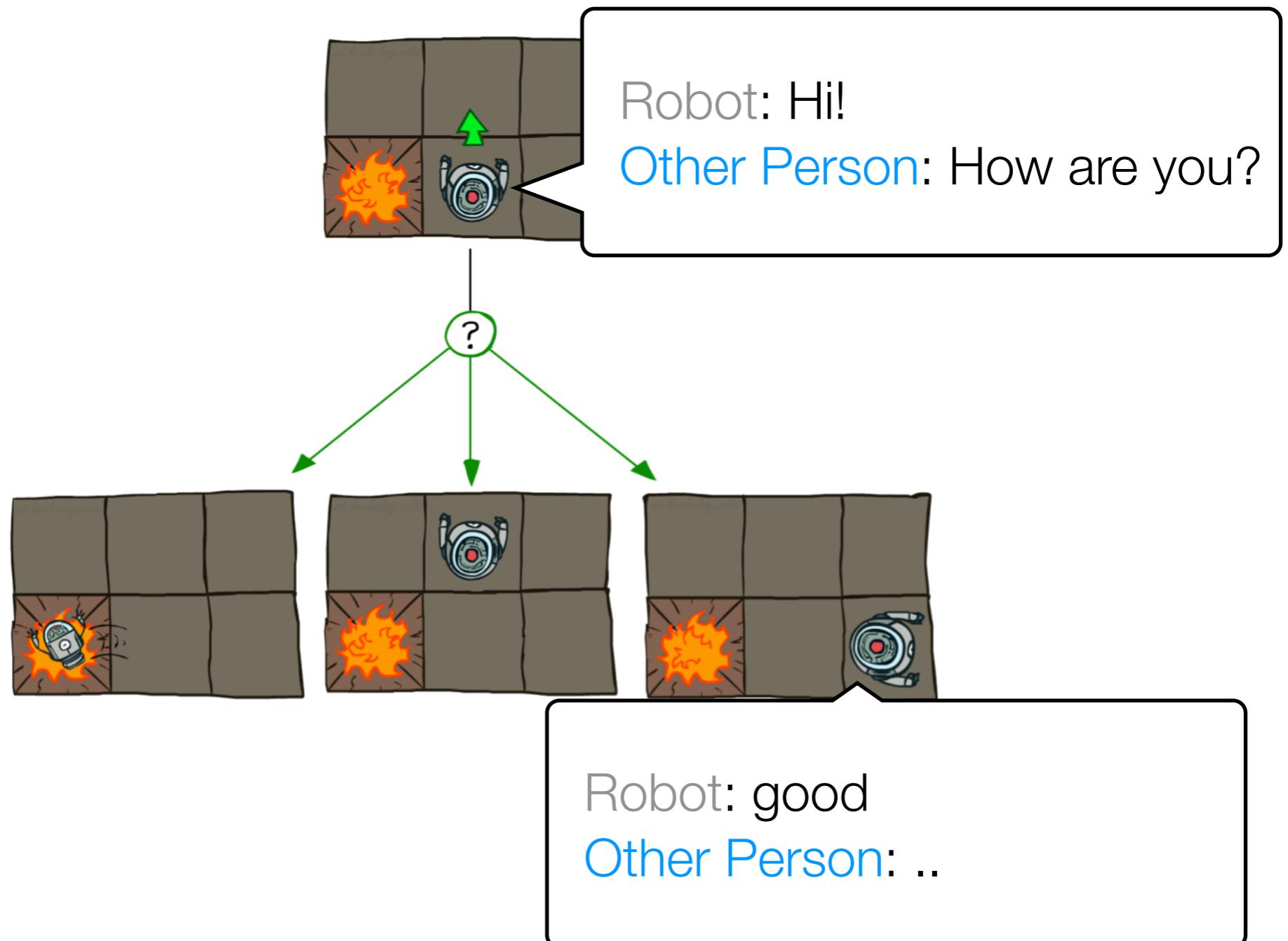
# Thinking of Reinforcement Learning for Dialog



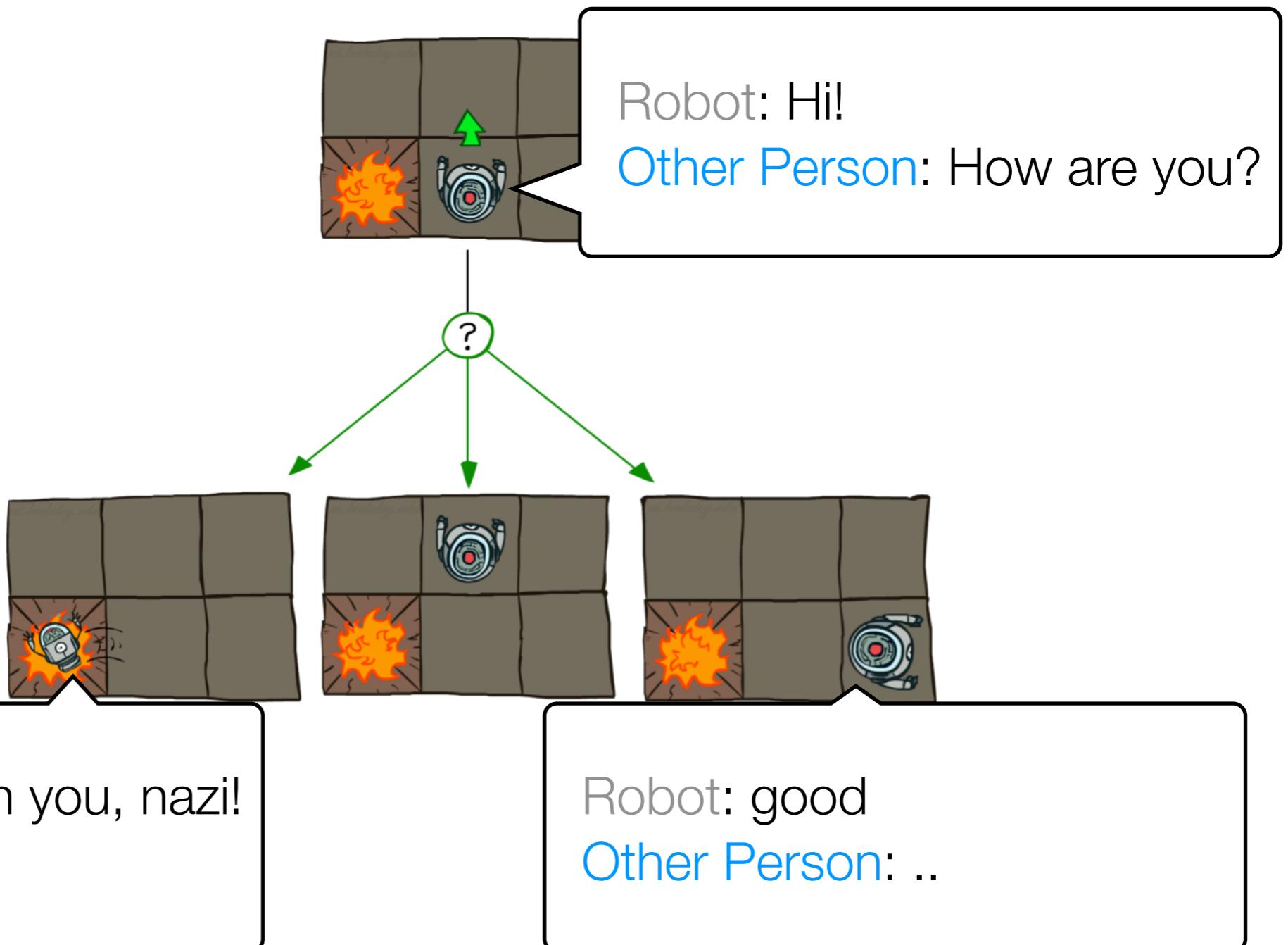
# Thinking of Reinforcement Learning for Dialog



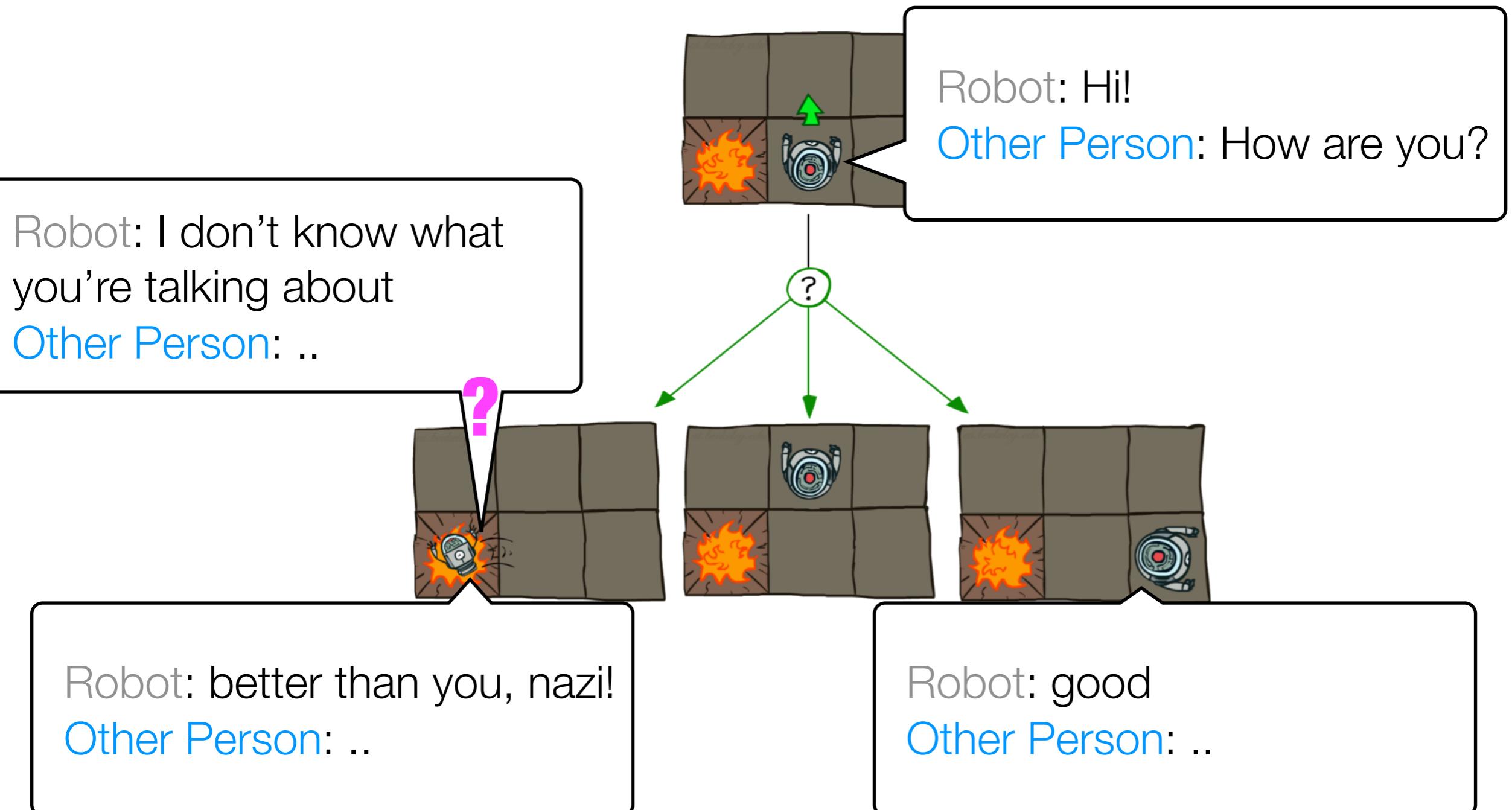
# Thinking of Reinforcement Learning for Dialog



# Thinking of Reinforcement Learning for Dialog



# Thinking of Reinforcement Learning for Dialog



Back to Reinforcement  
Learning for Dialog

# Issues

How do we handle long-term dialogue success?

- Problem 1: Dull and generic responses.

# Issues

- Problem 1: Dull and generic responses.

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015; )



Do you love me ?

I don't know what you are talking  
about.



# Issues

How do we handle long-term dialogue success?

- Problem 1: Dull and generic responses.
- Problem 2: Repetitive responses.

# Problem 2: Repetitive responses.



# Problem 2: Repetitive responses.

Shut up !



# Problem 2: Repetitive responses.

Shut up !

No, you shut  
up !



# Problem 2: Repetitive responses.

Shut up !

No, you shut  
up !

No, you shut up !



# Problem 2: Repetitive responses.

Shut up !

No, you shut  
up !

No, you shut up !

No, you shut  
up !



# Problem 2: Repetitive responses.



Shut up !

No, you shut up !

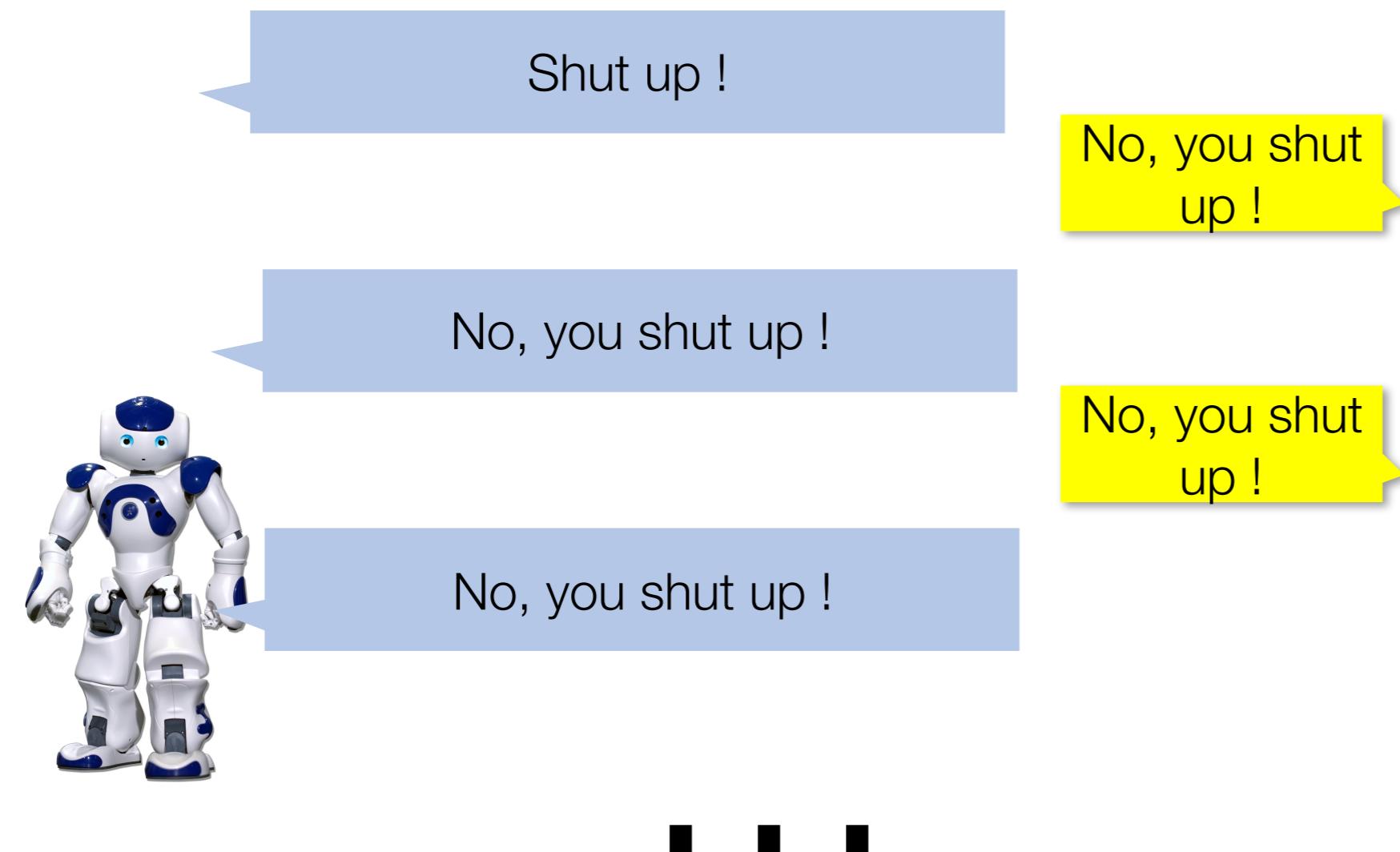
No, you shut up !

No, you shut  
up !

No, you shut  
up !



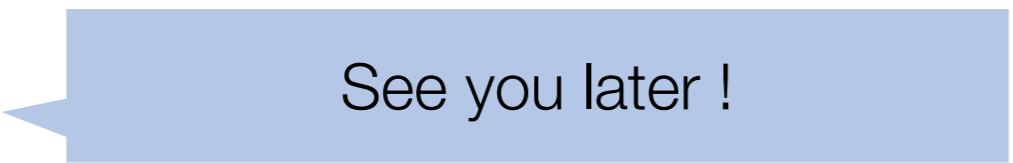
# Problem 2: Repetitive responses.



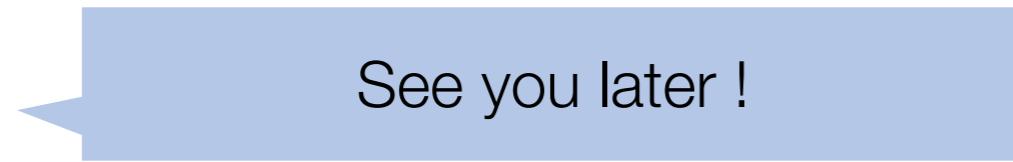
# Problem 2: Repetitive responses.



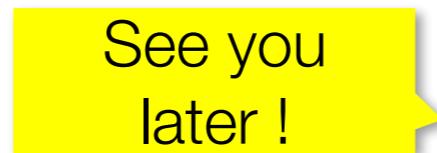
See you later !



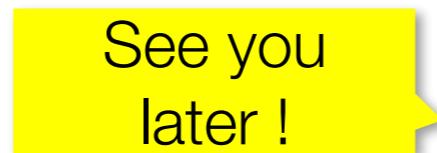
See you later !



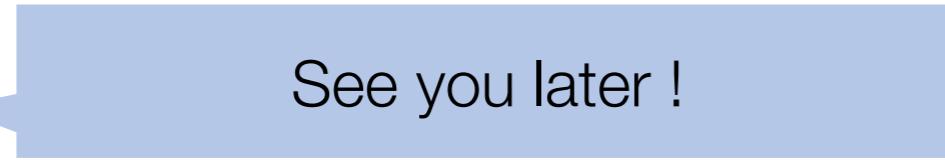
See you  
later !



See you  
later !



See you later !



■ ■ ■



# Issues

How do we handle long-term dialogue success?

- Problem 1: Dull and generic responses.
- Problem 2: Repetitive responses.
- Problem 3: Short-sighted conversation decisions.

# Problem 3: Short-sighted conversation decisions.



# Problem 3: Short-sighted conversation decisions.

How old are you ?



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



# Problem 3: Short-sighted conversation decisions.



How old are you ?

i 'm 16 .

16 ?



# Problem 3: Short-sighted conversation decisions.



How old are you ?

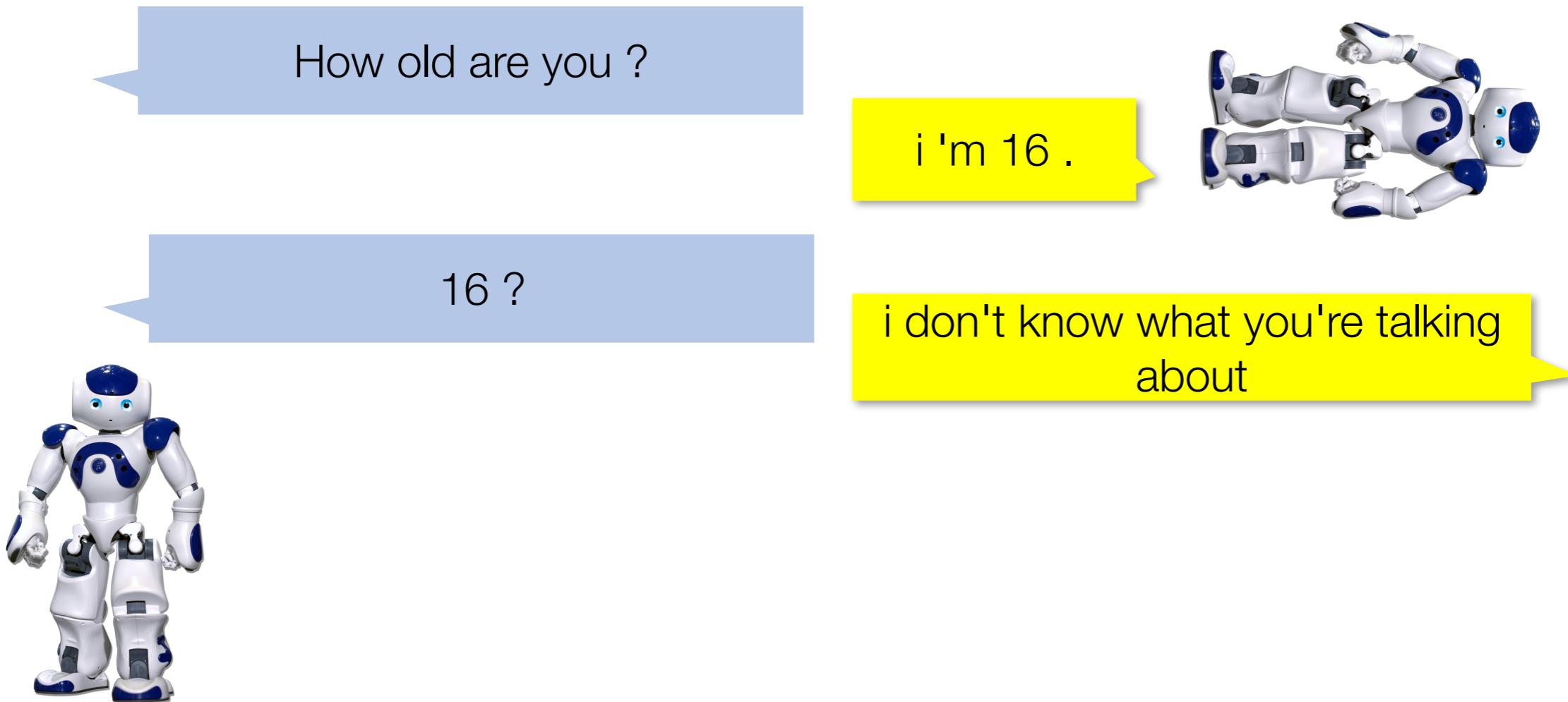
i 'm 16 .

16 ?

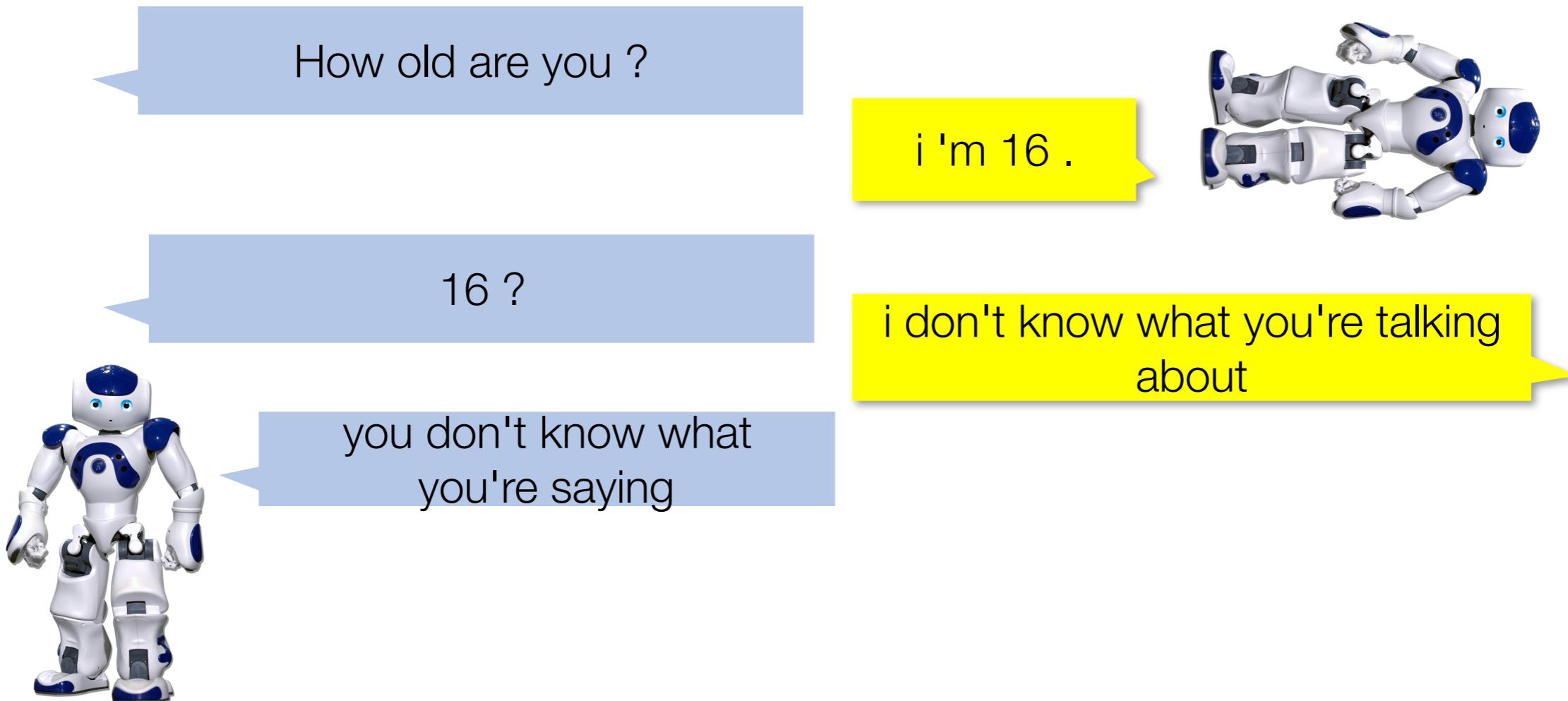


i don't know what you're talking  
about

# Problem 3: Short-sighted conversation decisions.



# Problem 3: Short-sighted conversation decisions.



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don't know what you're talking  
about

you don't know what  
you're saying



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don't know what you're talking  
about

you don't know what  
you're saying

i don't know what you're talking  
about



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don't know what you're talking  
about

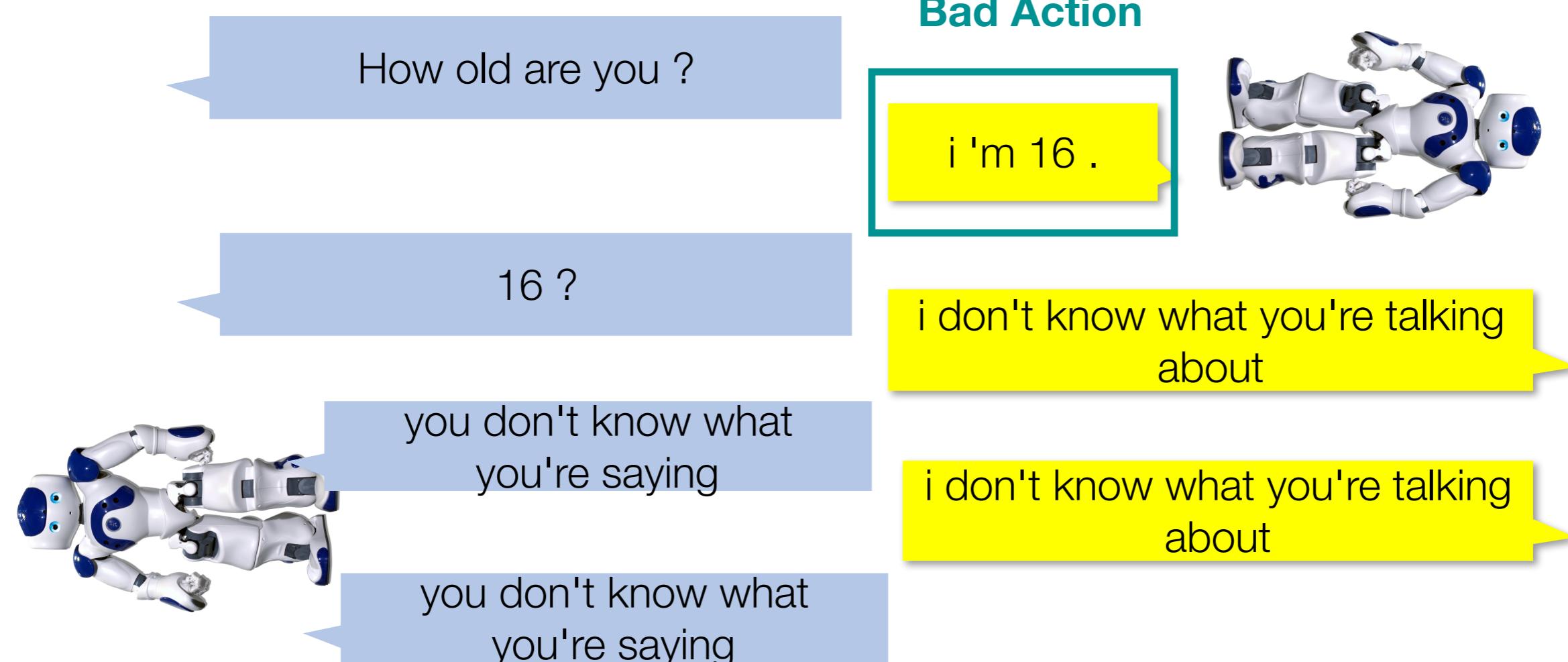
you don't know what  
you're saying

i don't know what you're talking  
about

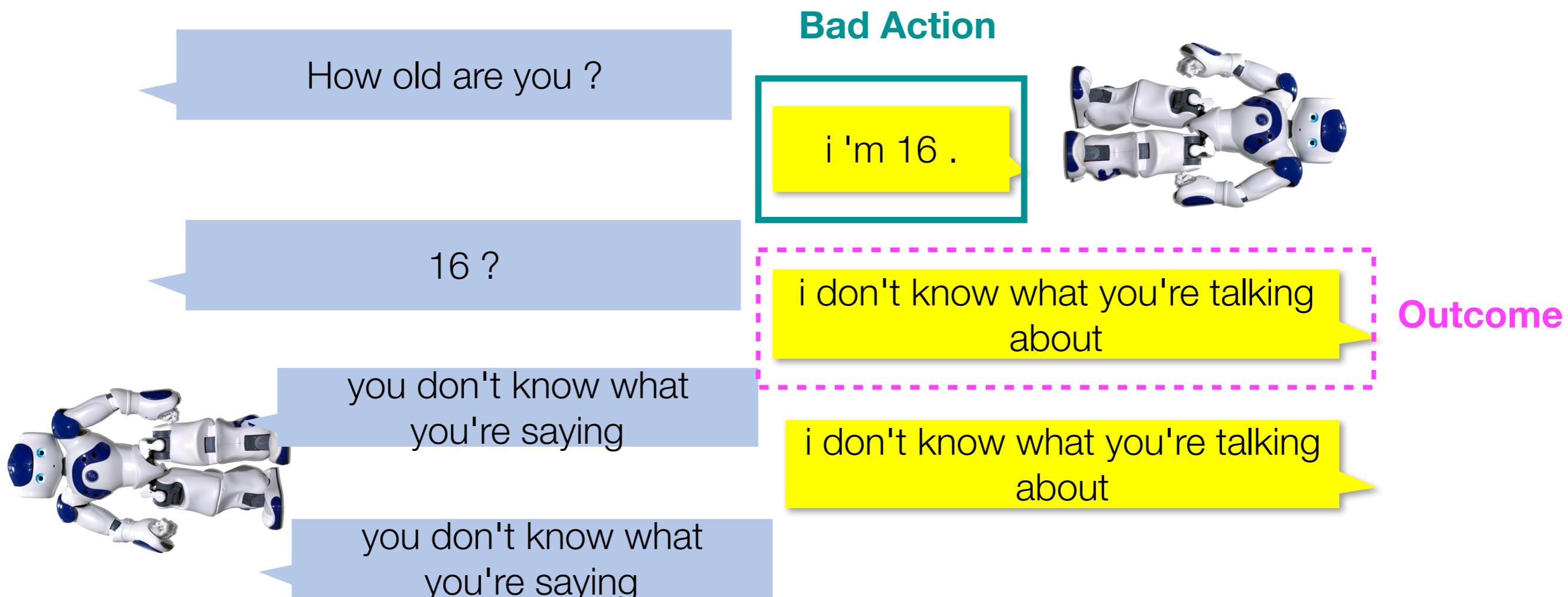


you don't know what  
you're saying

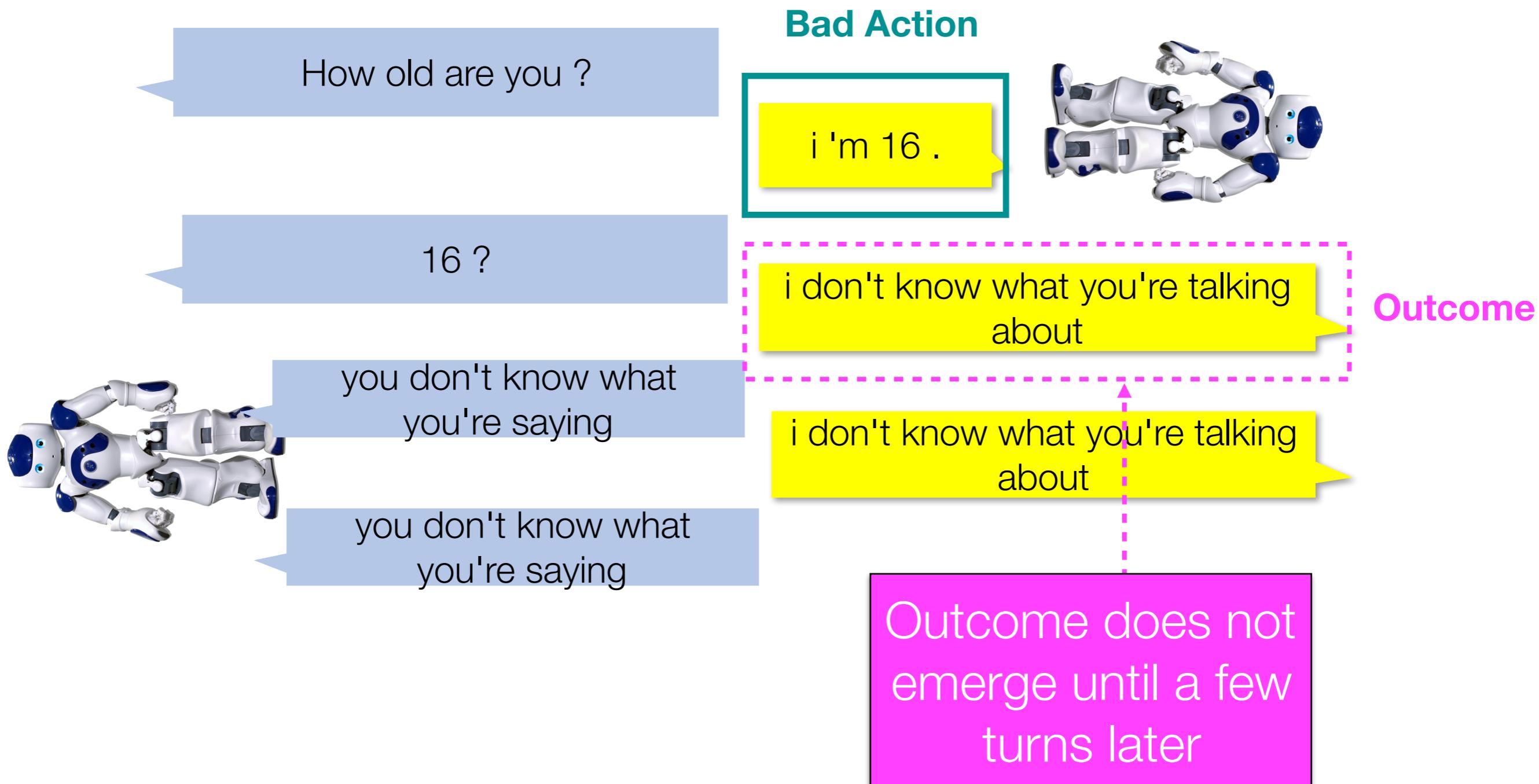
# Problem 3: Short-sighted conversation decisions.



# Problem 3: Short-sighted conversation decisions.

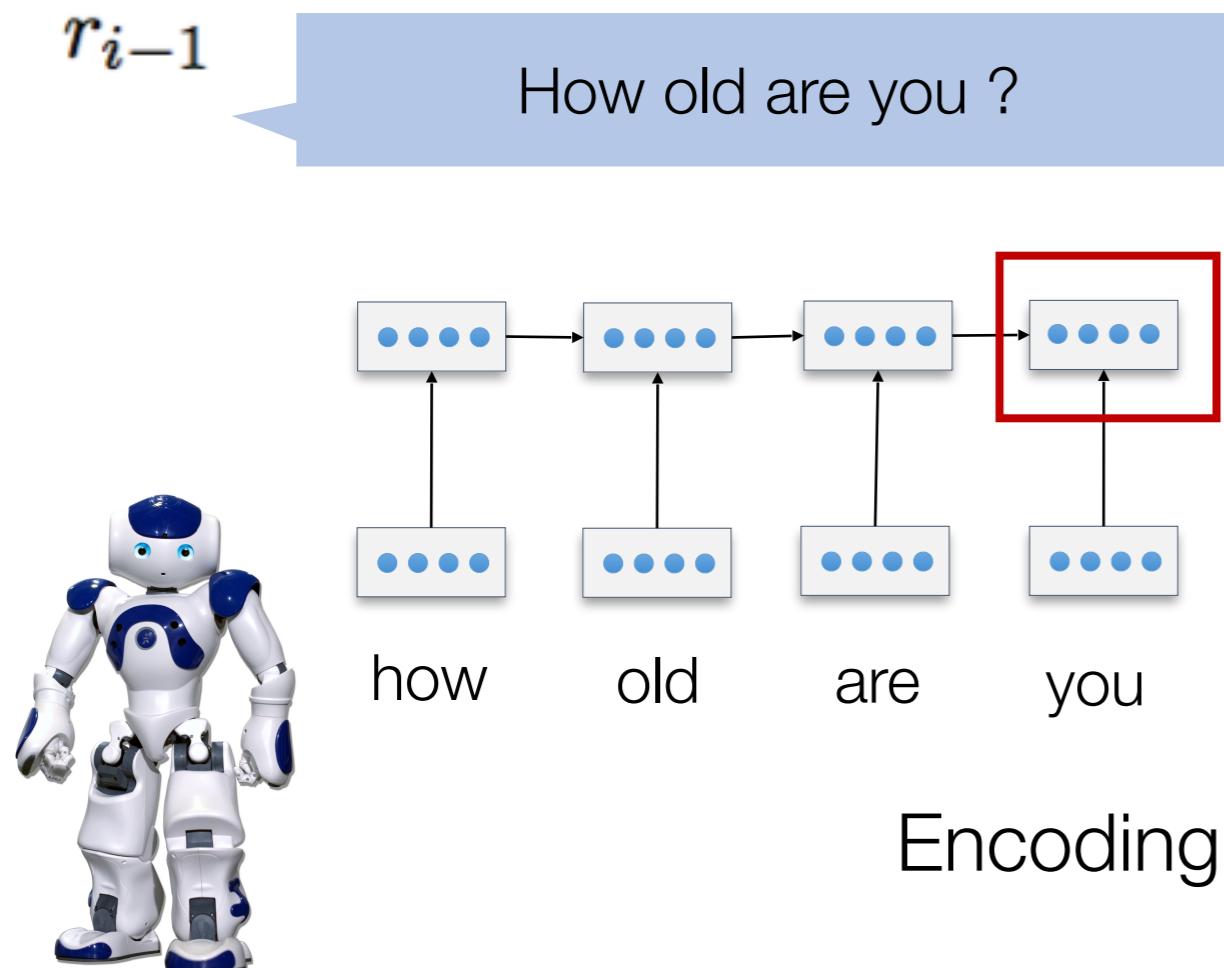


# Can reinforcement learning handle this?



# Notations for Reinforcement Learning

# Notations: State



# Notations: Reward

$r_{i-1}$

How old are you ?



$r_i$

i 'm 16 .



# Notations: Reward

$r_{i-1}$

How old are you ?

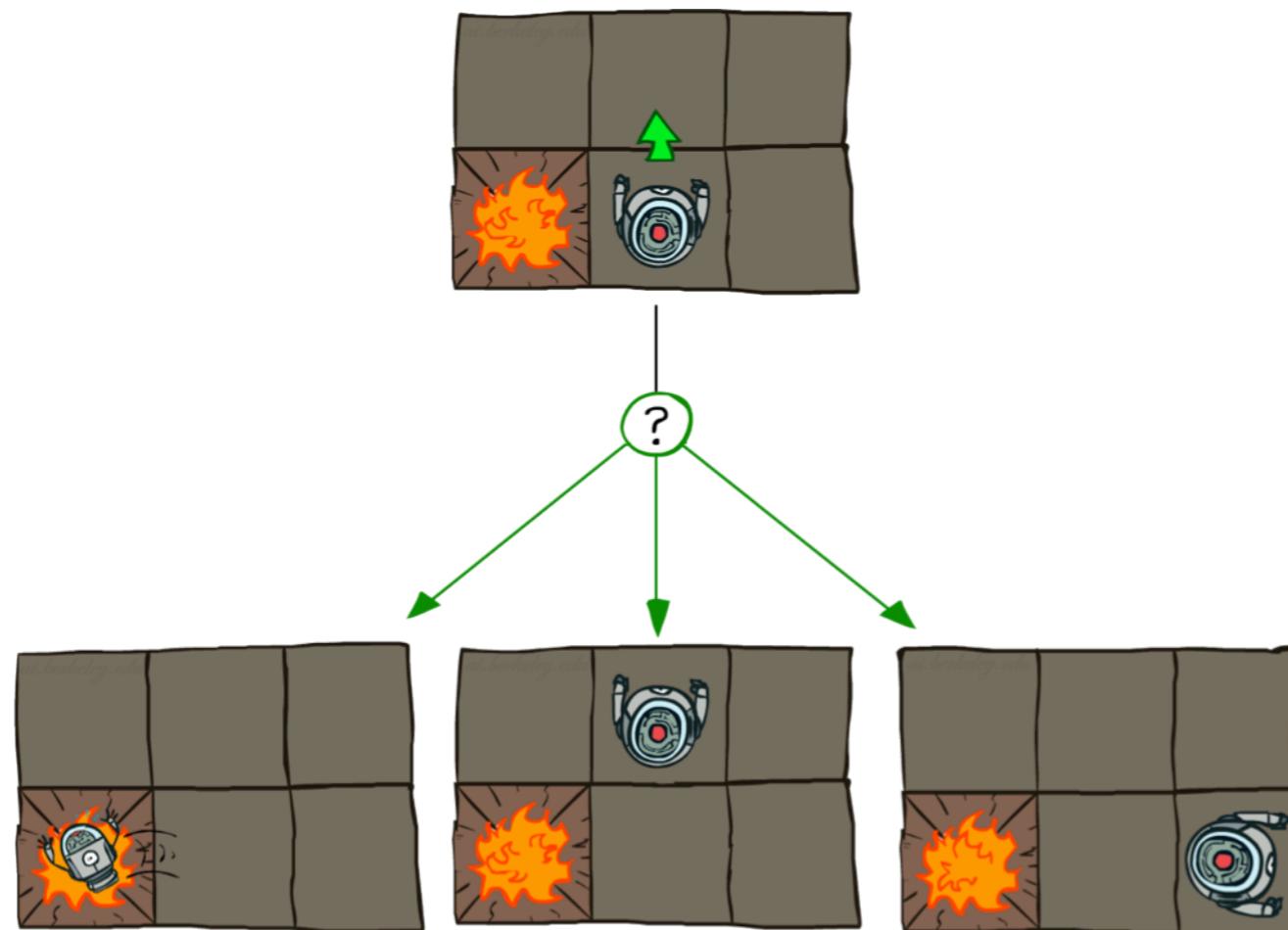


$r_i$

i'm 16 .

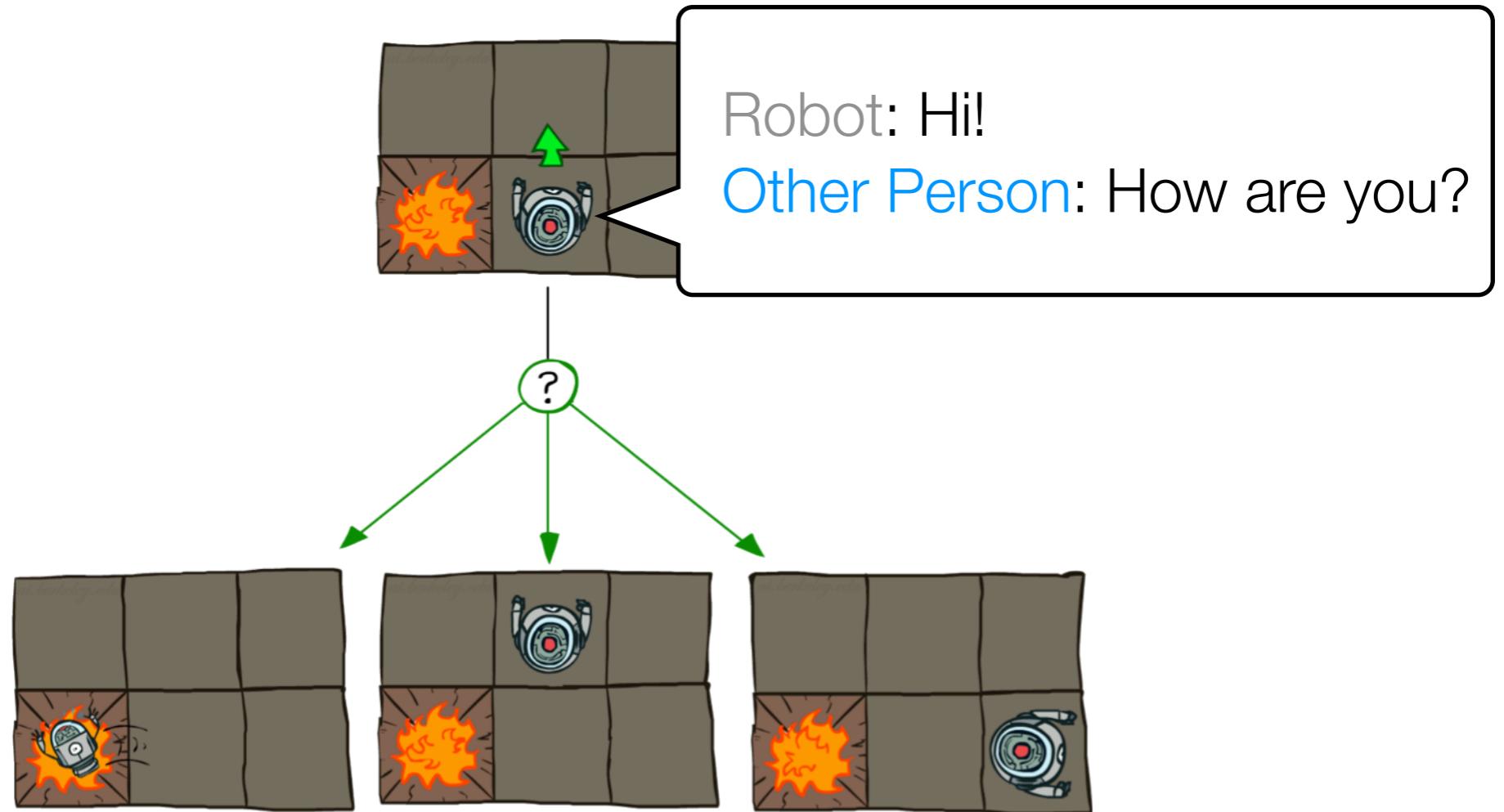


# Breakout Session Time



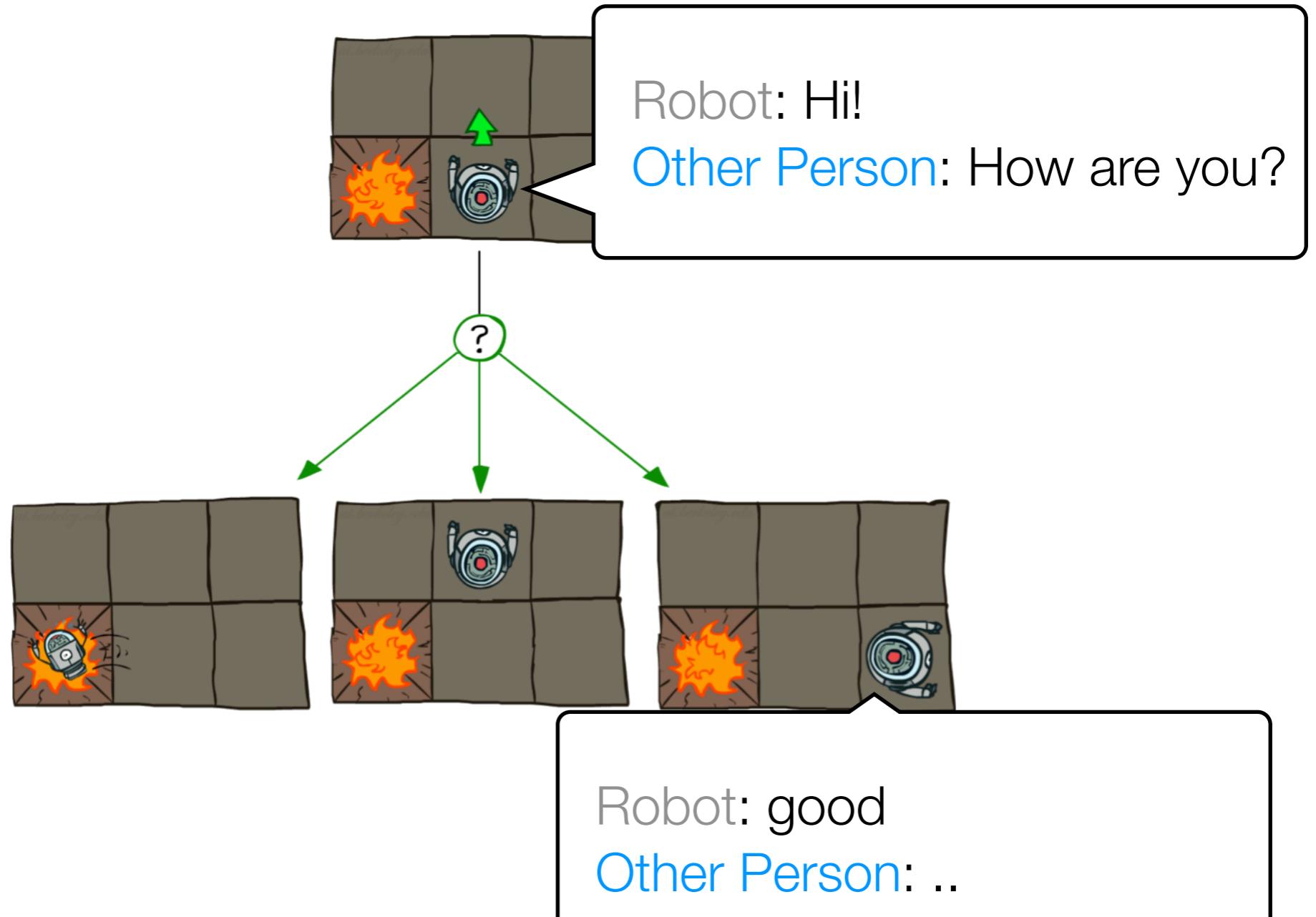
What features of a conversation would make for a good reward?  
→ What can we computationally measure that we would want to maximize or minimize? (5 min)

# Breakout Session Time



What features of a conversation would make for a good reward?  
→ What can we computationally measure that we would want to maximize or minimize? (5 min)

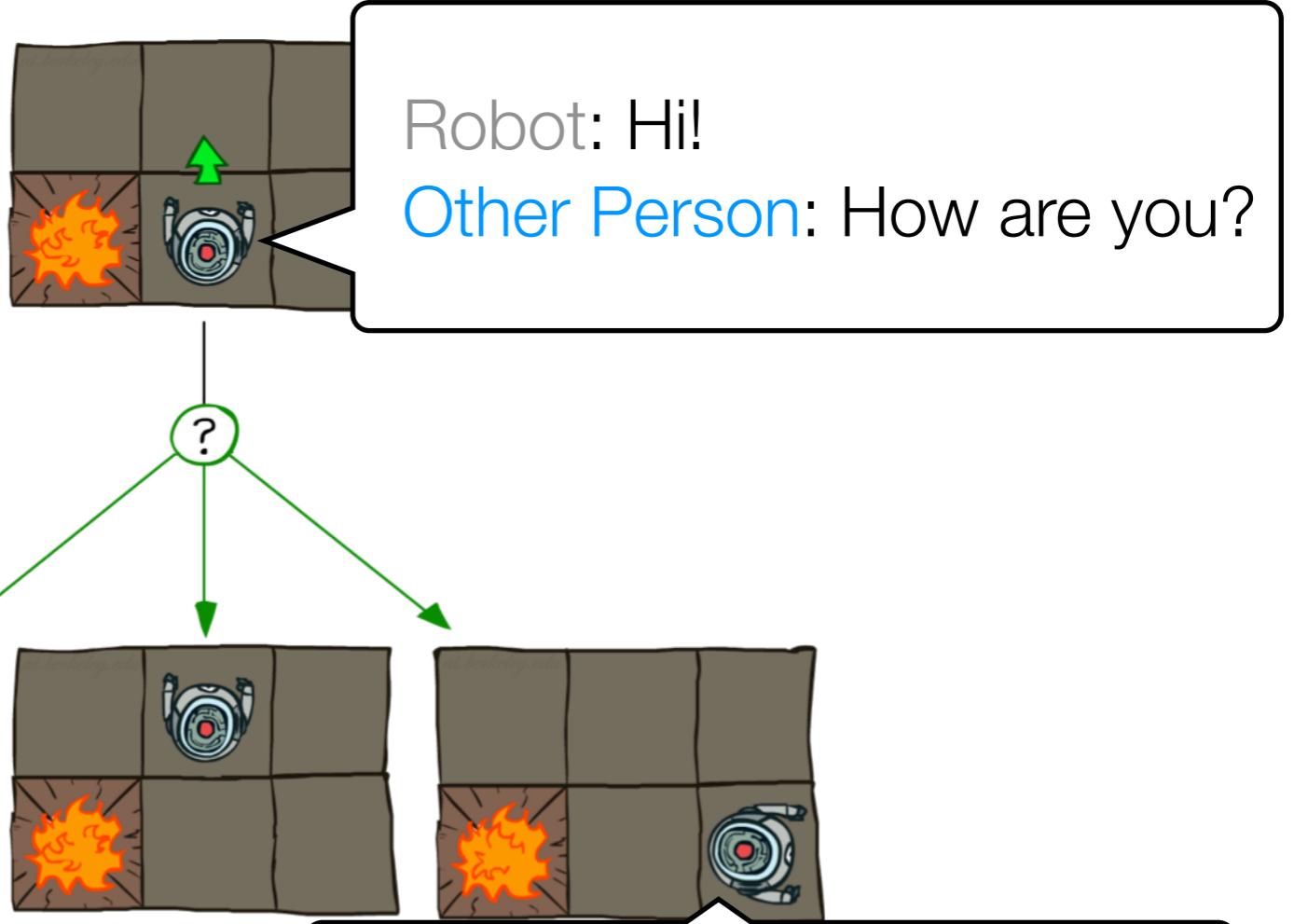
# Breakout Session Time



What features of a conversation would make for a good reward?  
→ What can we computationally measure that we would want to maximize or minimize? (5 min)

# Breakout Session Time

Robot: I don't know what you're talking about  
Other Person: ...



Robot: better than you, nazi!  
Other Person: ...

Robot: good  
Other Person: ...

What features of a conversation would make for a good reward?  
→ What can we computationally measure that we would want to maximize or minimize? (5 min)

# Notations: Reward

1. Ease of answering

$$R_1 = - \sum_{s \in \mathbb{S}} \log p_{\text{seq2seq}}(s|a)$$

# Notations: Reward

## 1. Ease of answering

$$R_1 = - \sum_{s \in \mathbb{S}} \log p_{\text{seq2seq}}(s|a)$$

S: "I don't know what you are talking about"

# Notations: Reward

## 2. Information Flow



See you later !



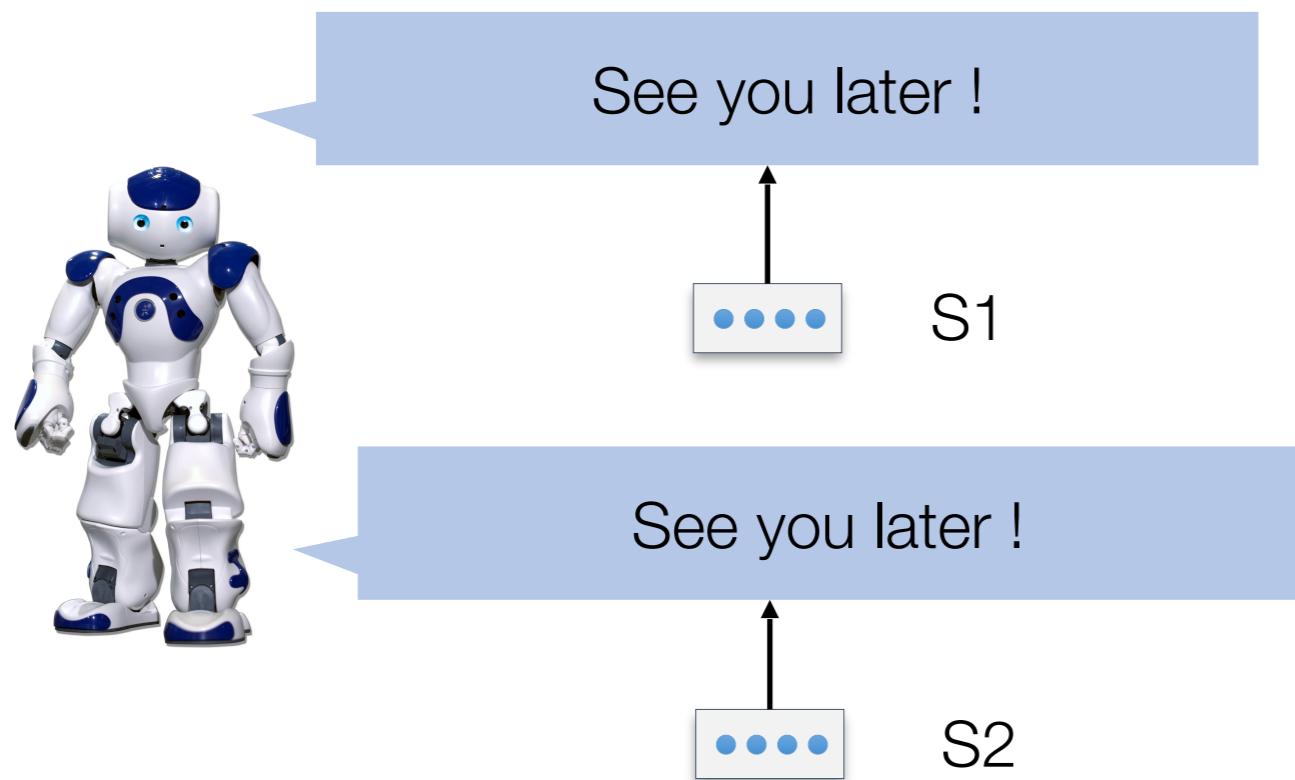
See you  
later !

See you later !

See you  
later !

# Notations: Reward

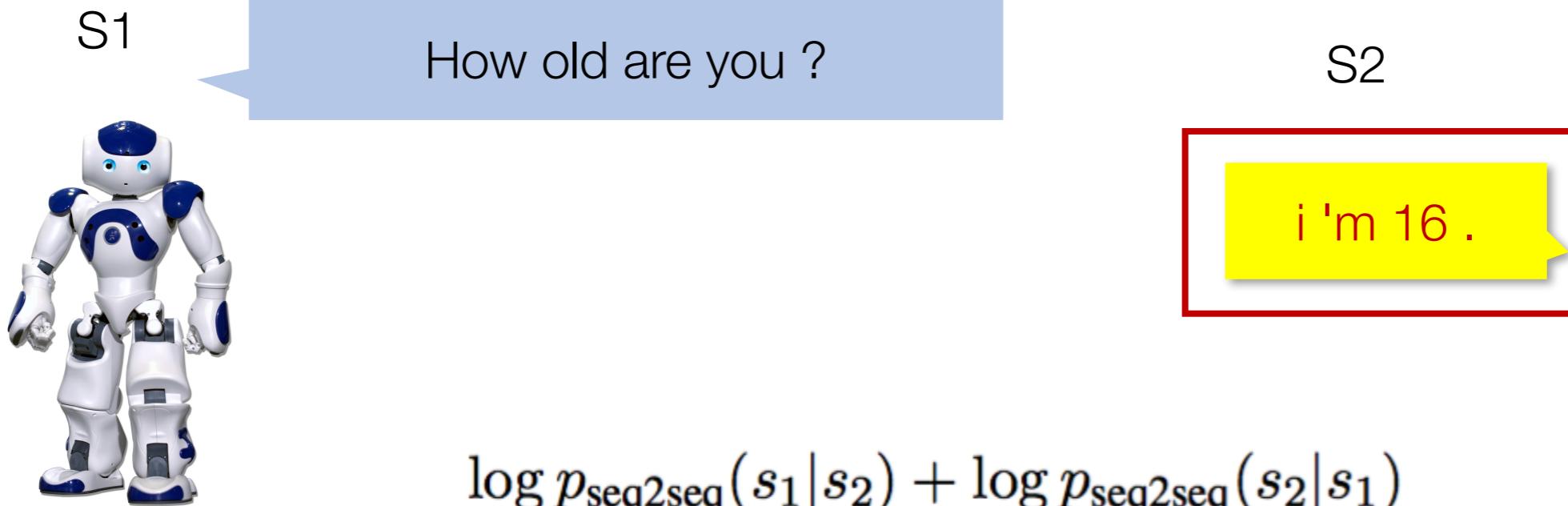
## 2. Information Flow



$$r_2 = -\log \text{Sigmoid}(\cos(s_1, s_2))$$

# Notations: Reward

## 3. Meaningfulness



# Notations: Reward

1. Easy to answer      R1
2. Information Flow    R2
3. Meaningfulness     R3

$$R = \lambda_1 R_1 + \lambda_2 R_2 + \lambda_3 R_3$$

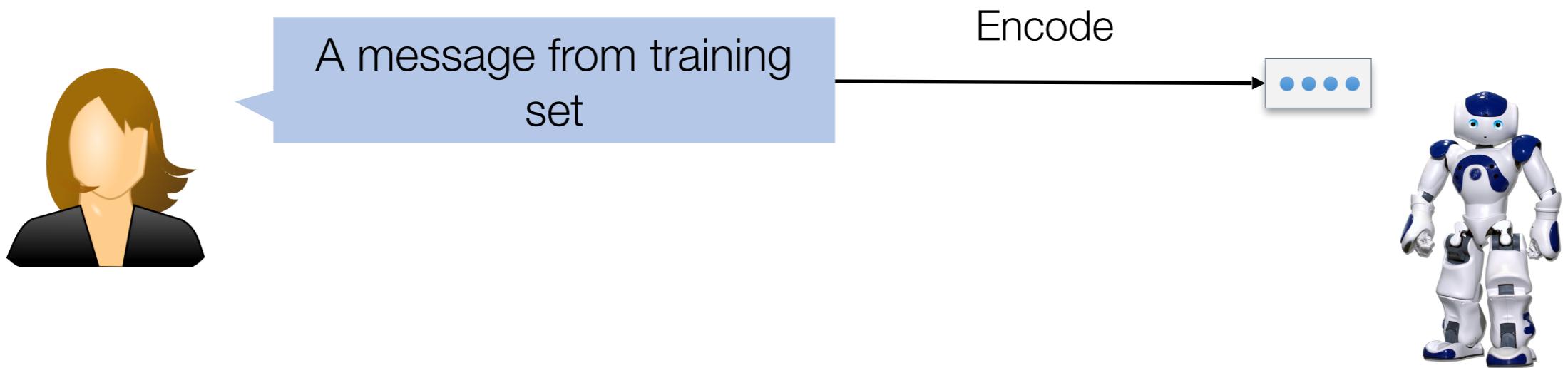
# Successful dialogue

- Ease of answering. A dialogue turn should be easy to response to. Operationalize: negative log likelihood of a “dull” response (“I don’t know what you’re talking about”; “I have no idea”).
- Information flow. Turns should add new information. Operationalize: negative log of cosine similarity between turns
- Semantic coherence: Turns should make sentence given the previous turns.

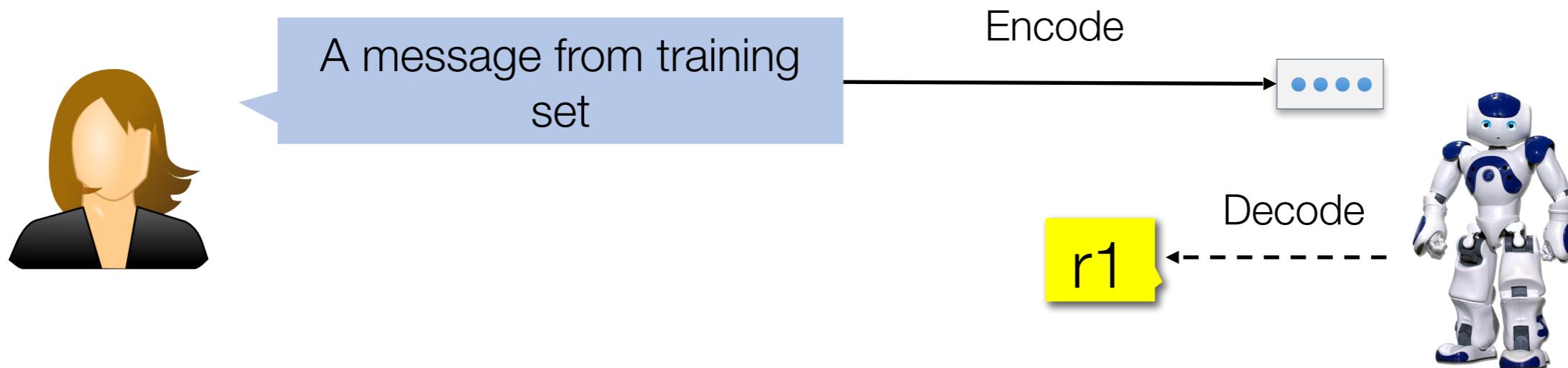
# Simulation



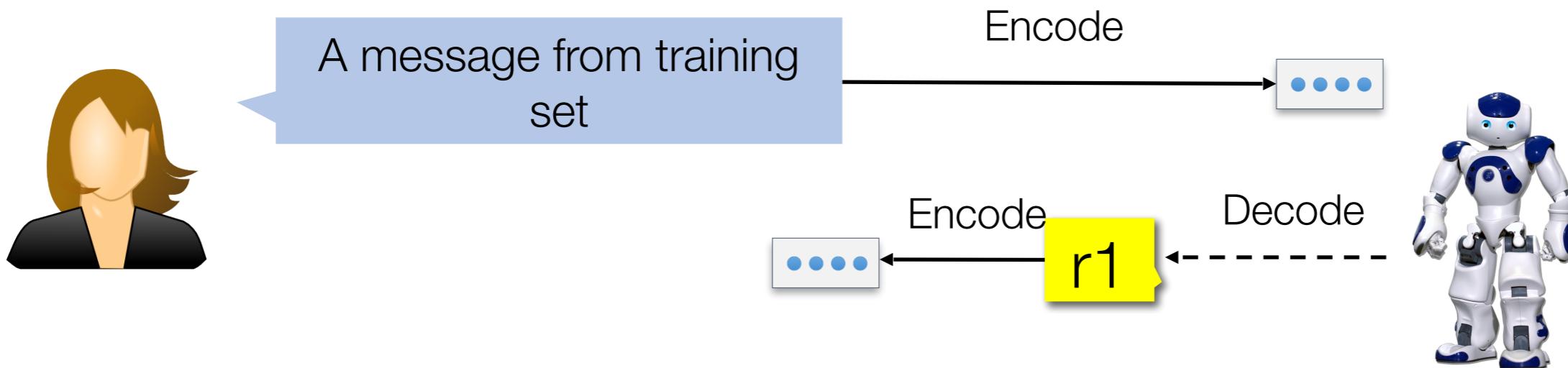
# Simulation



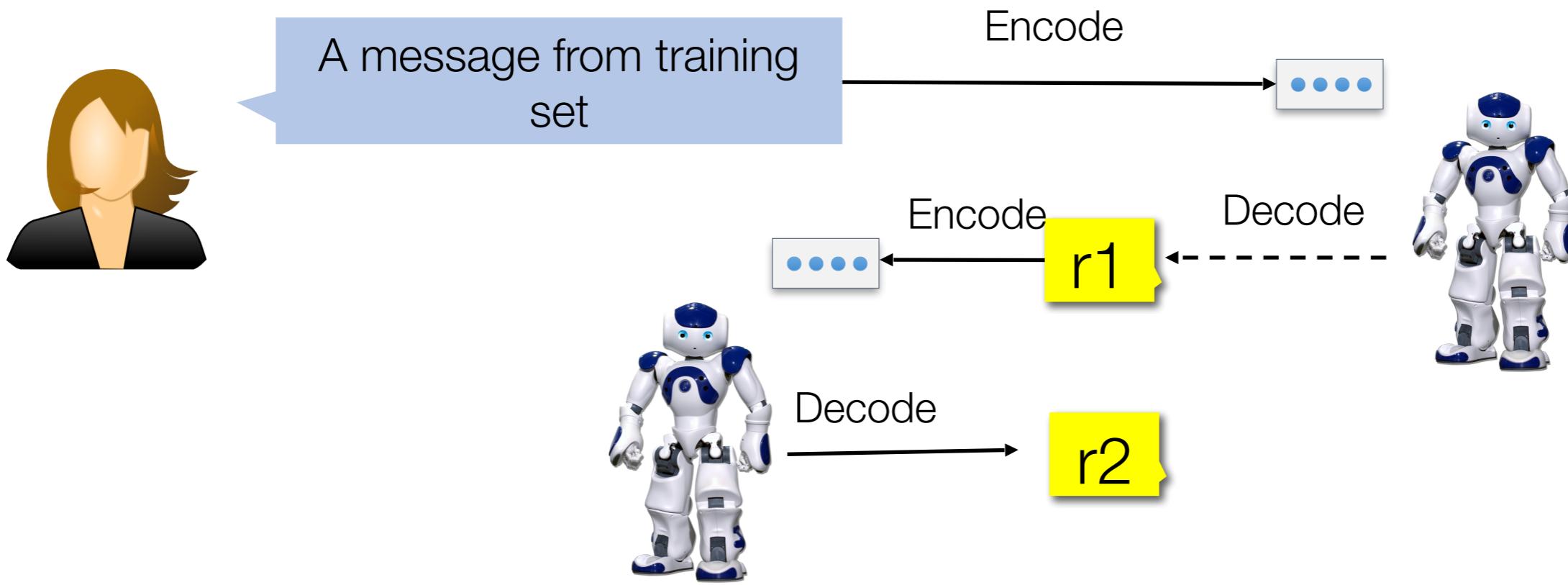
# Simulation



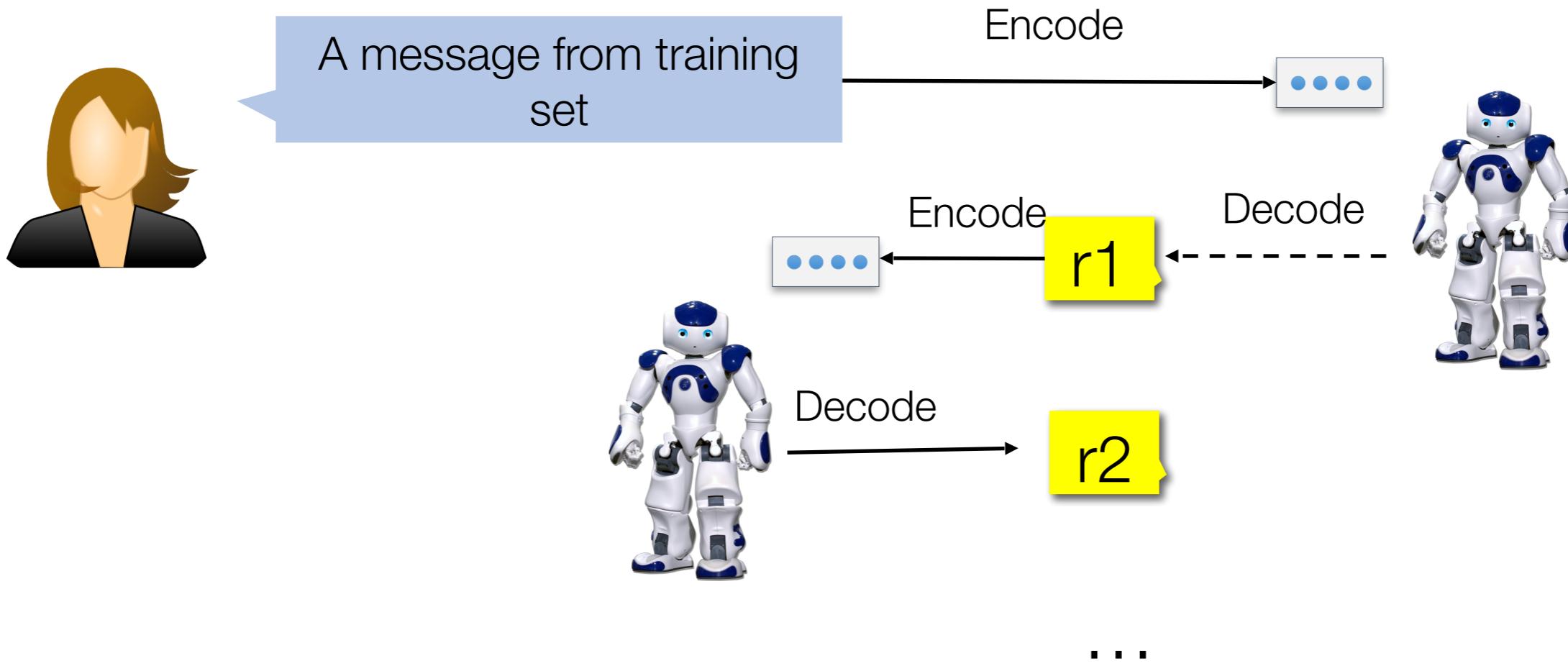
# Simulation

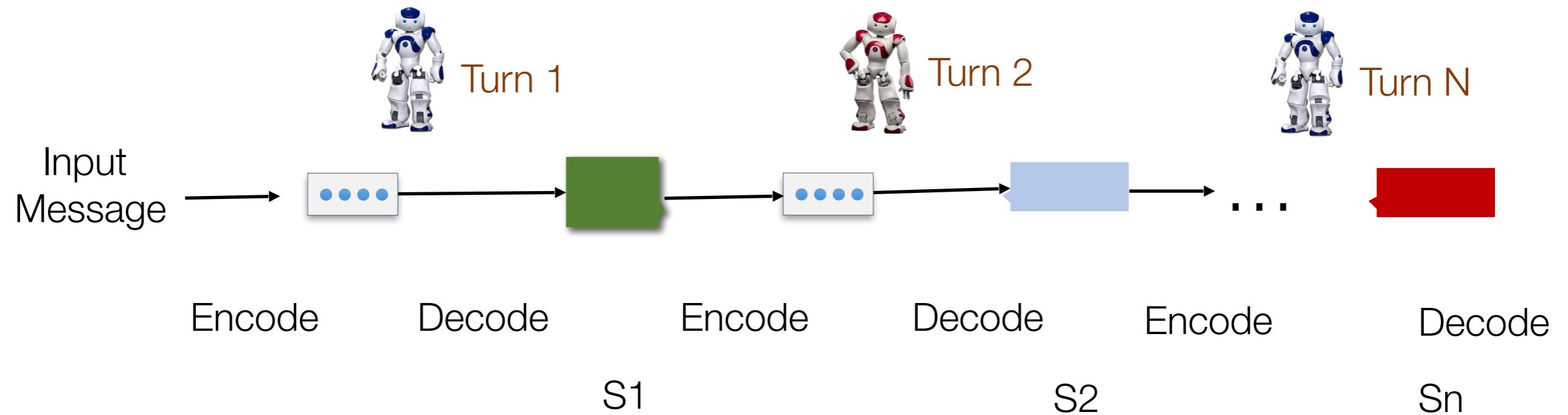


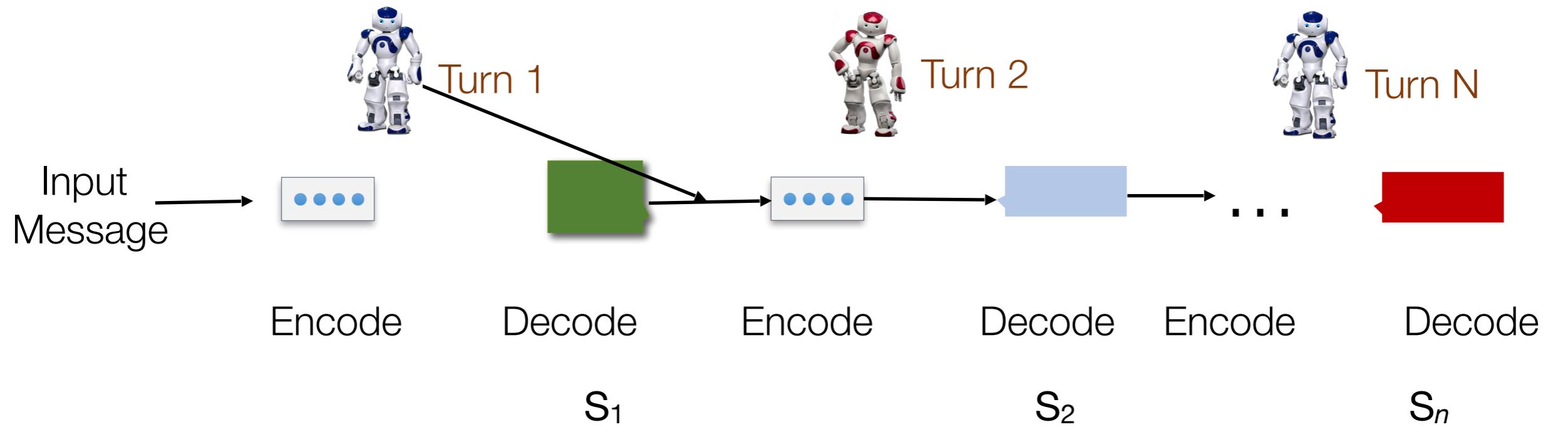
# Simulation



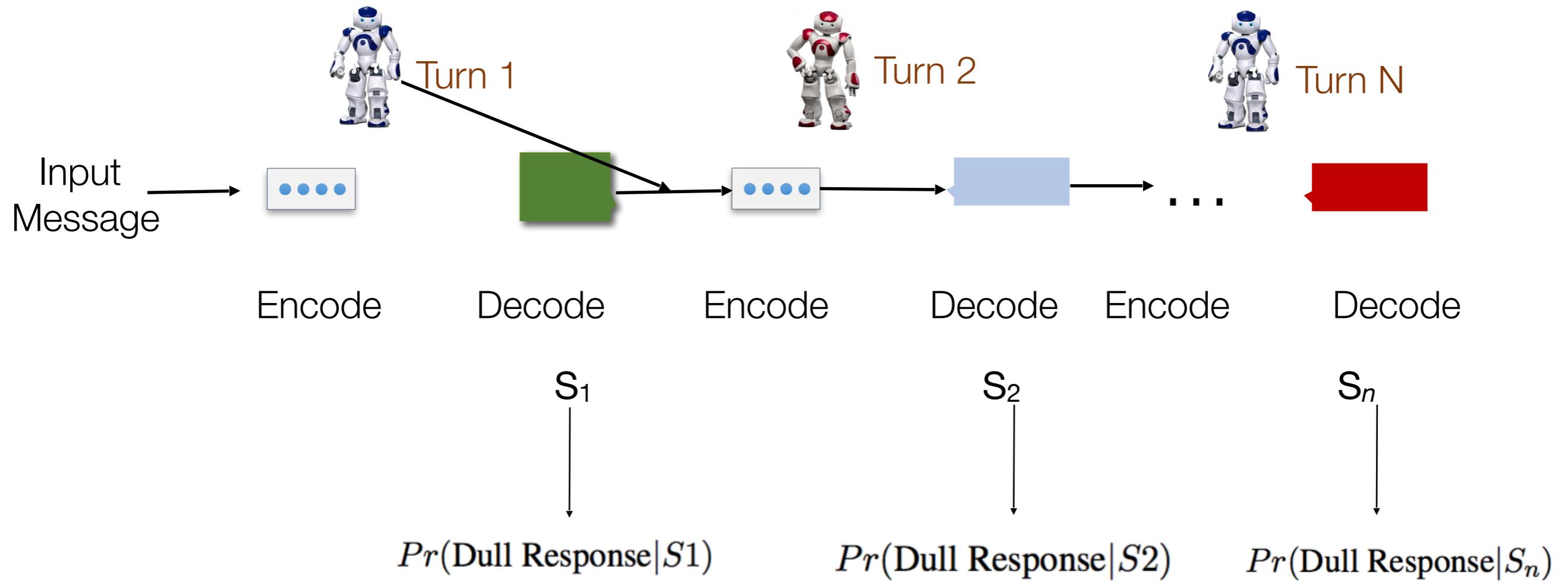
# Simulation



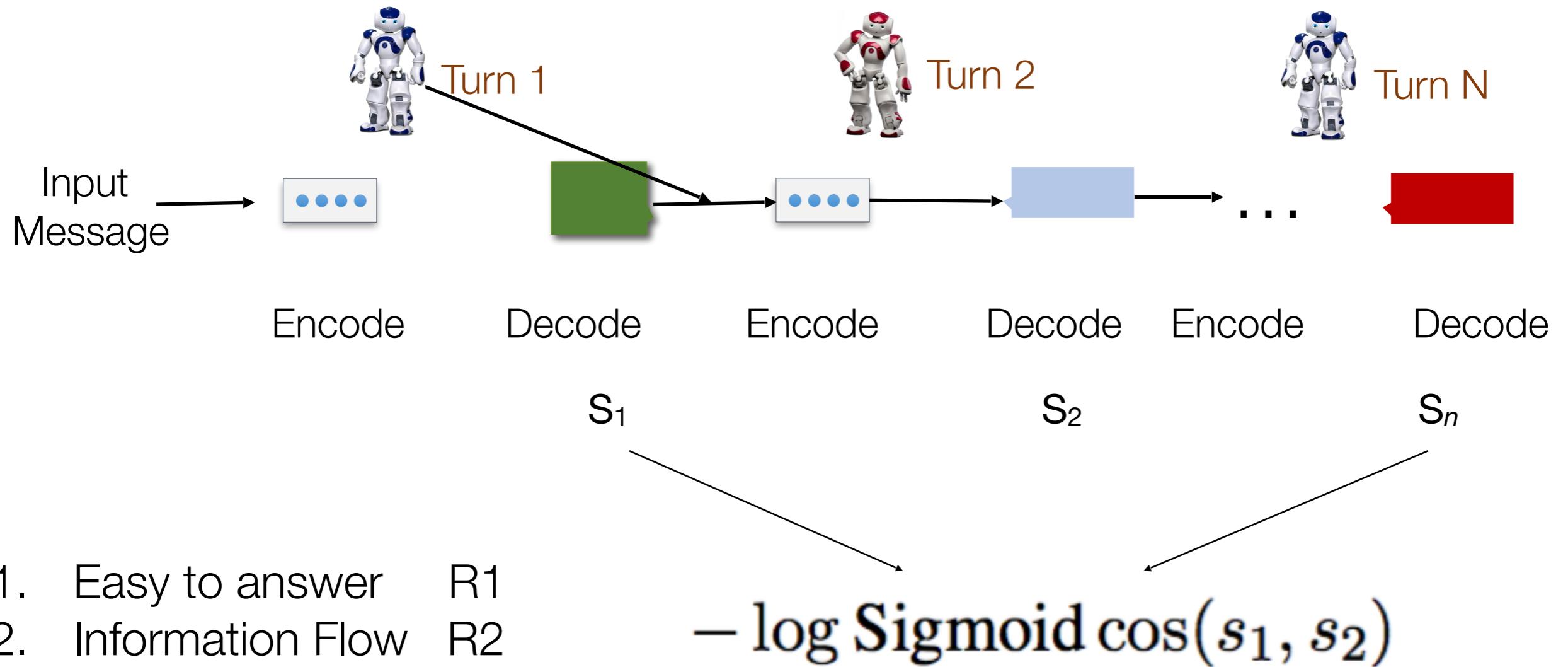


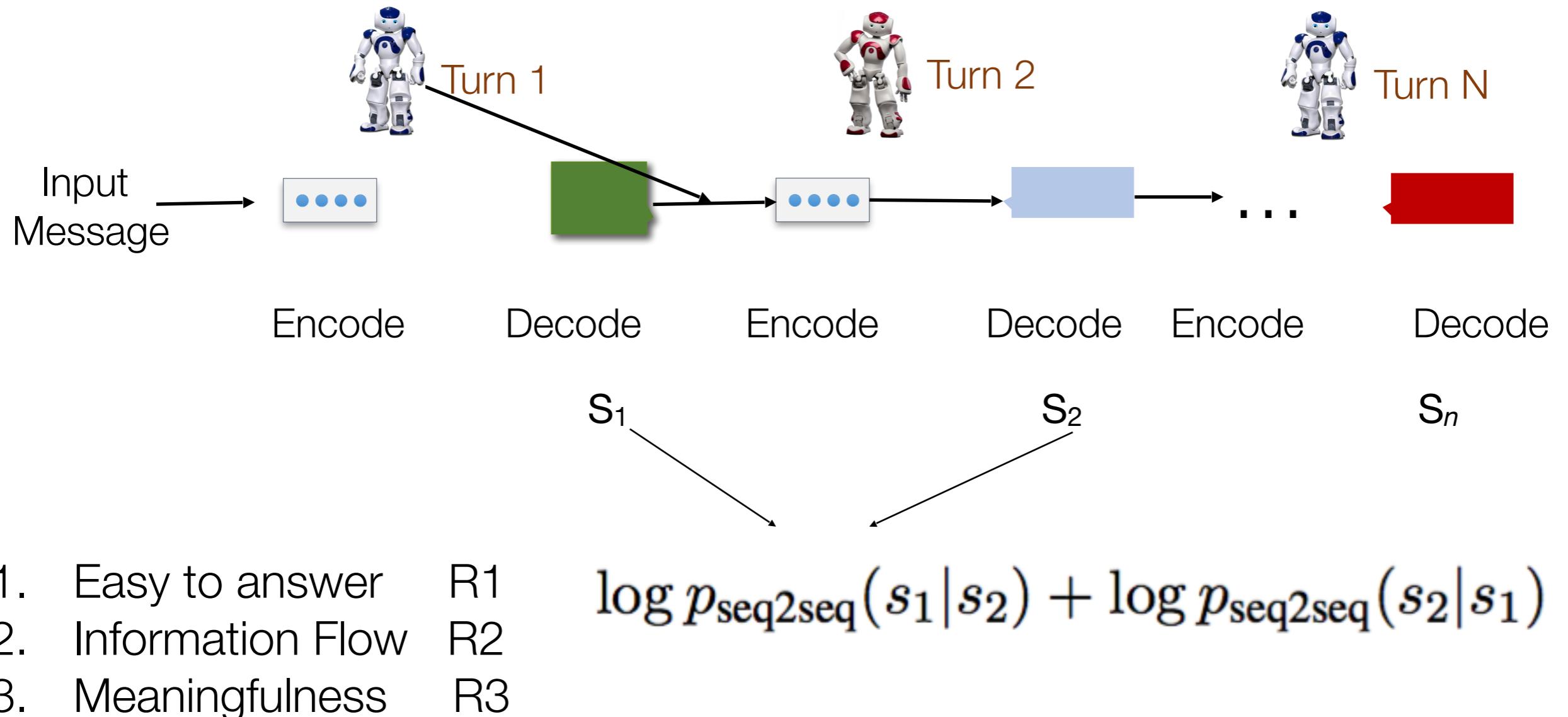


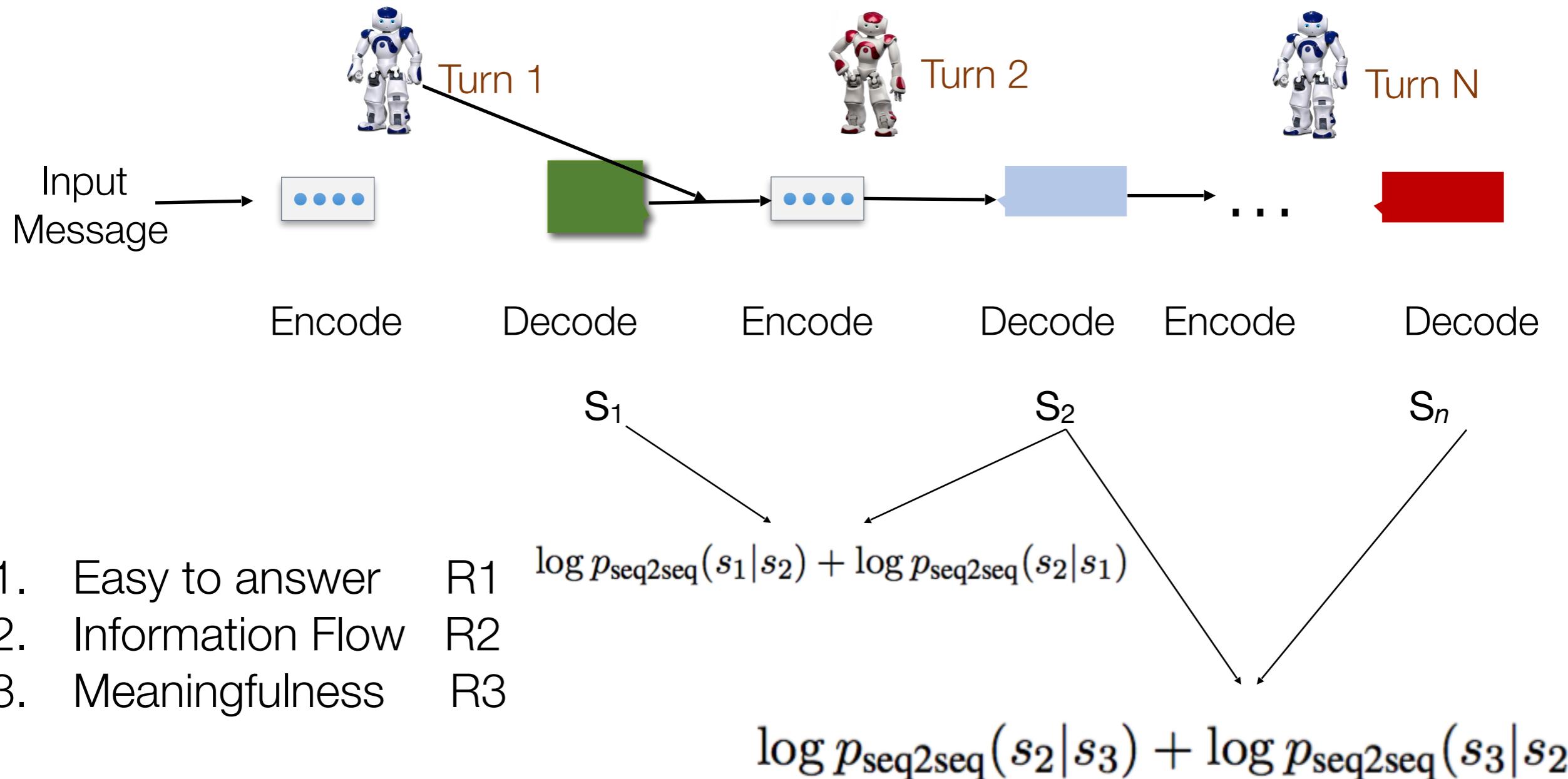
Compute Accumulated Reward  $R(S_1, S_2, \dots, S_n)$

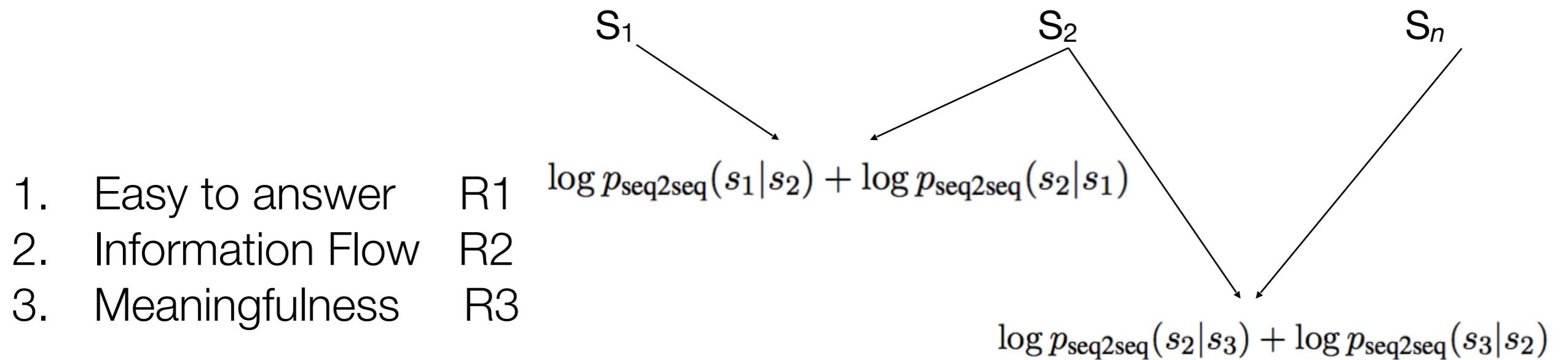
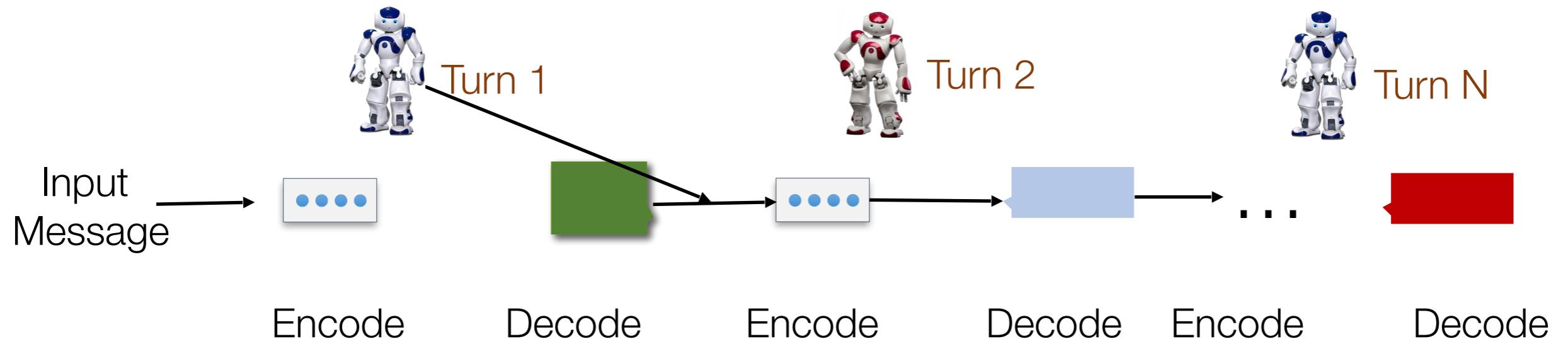


1. Easy to answer R1

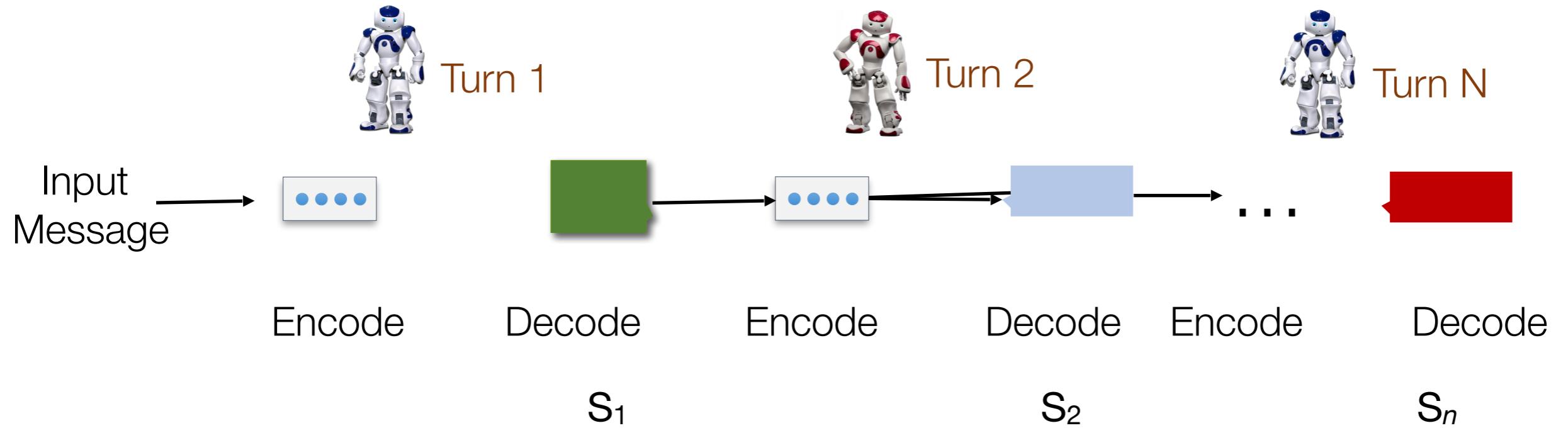






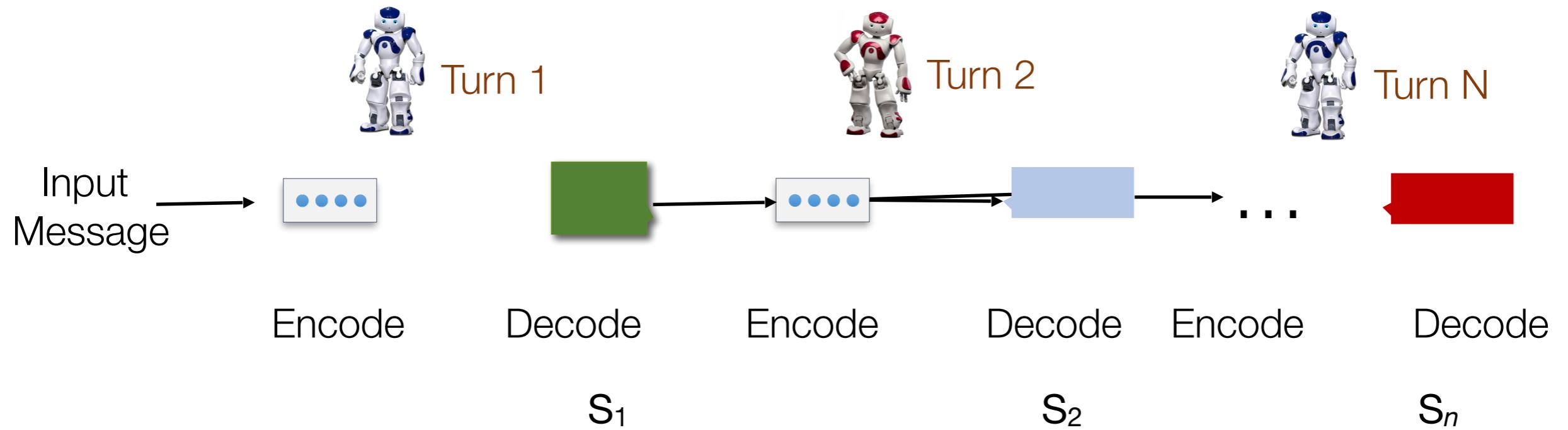


Compute Accumulated Reward  $R(S_1, S_2, \dots, S_n)$



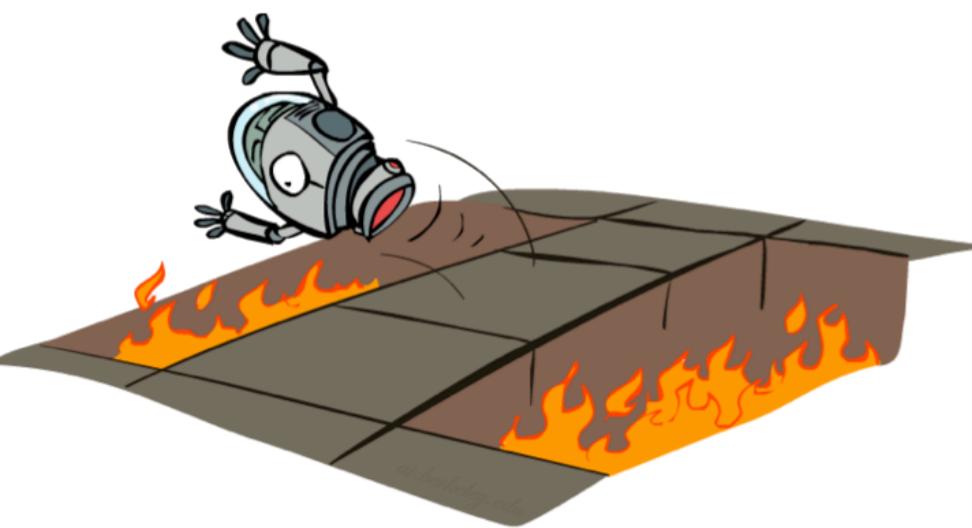
REINFORCE Algorithm  
(William, 1992)

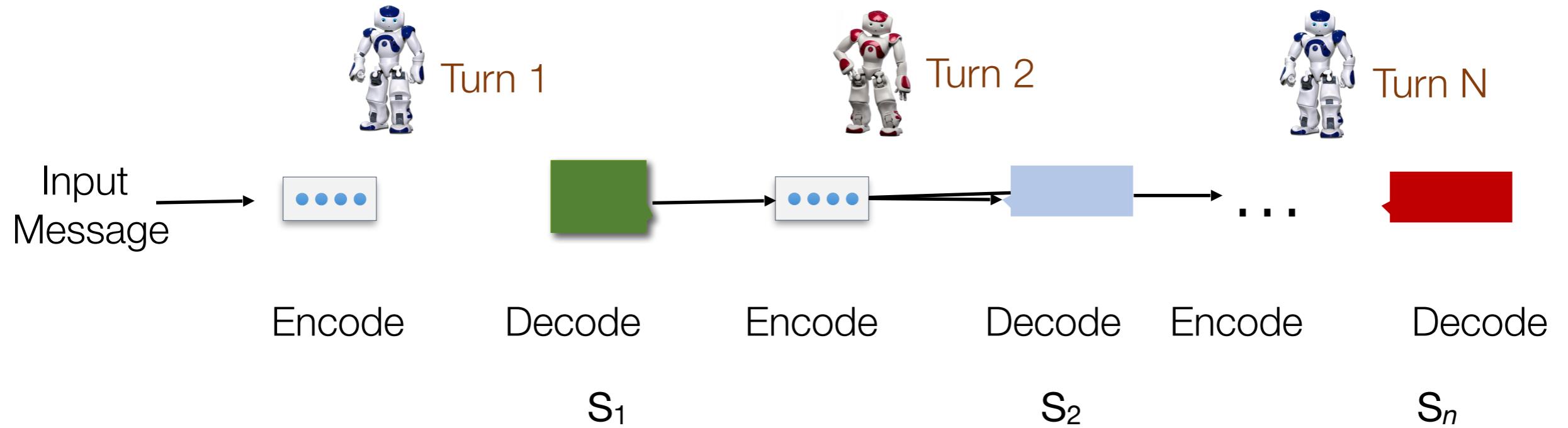
$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$



REINFORCE Algorithm  
(William, 1992)

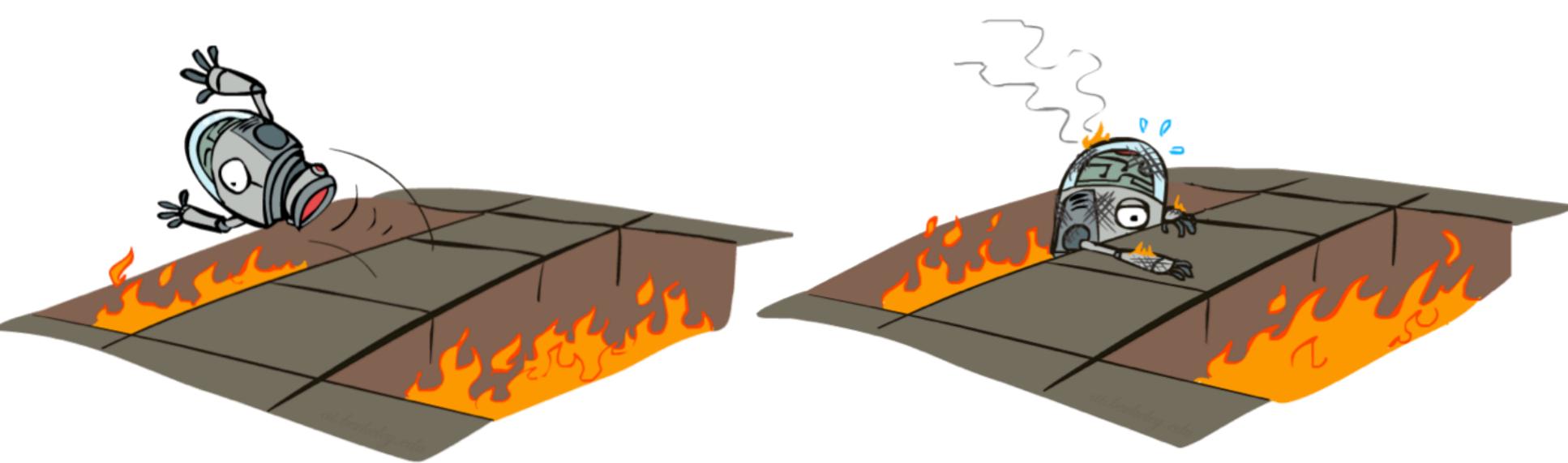
$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

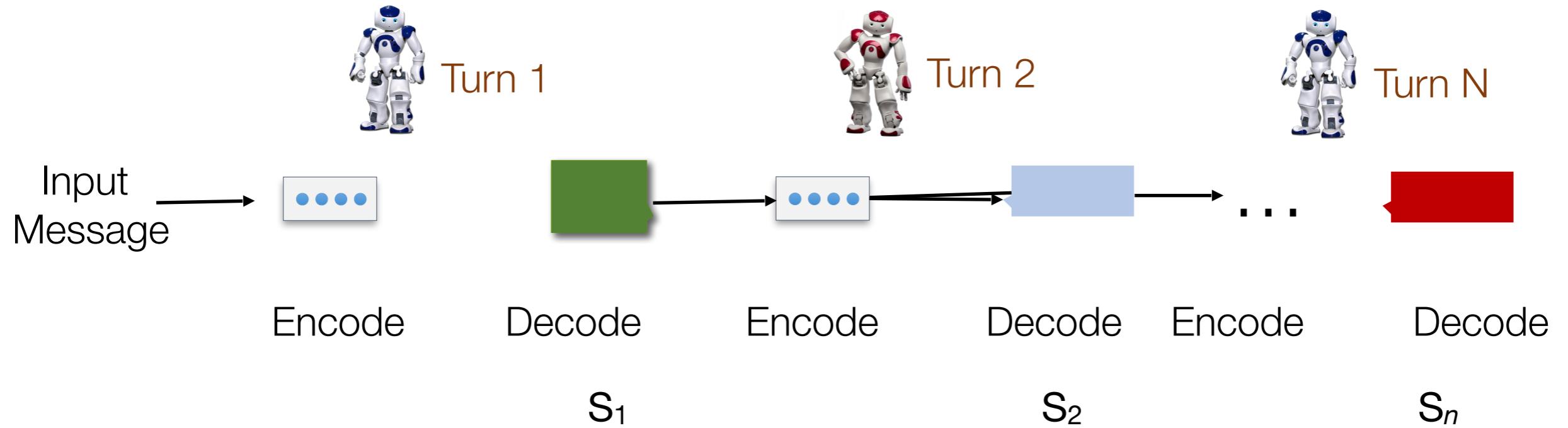




REINFORCE Algorithm  
(William, 1992)

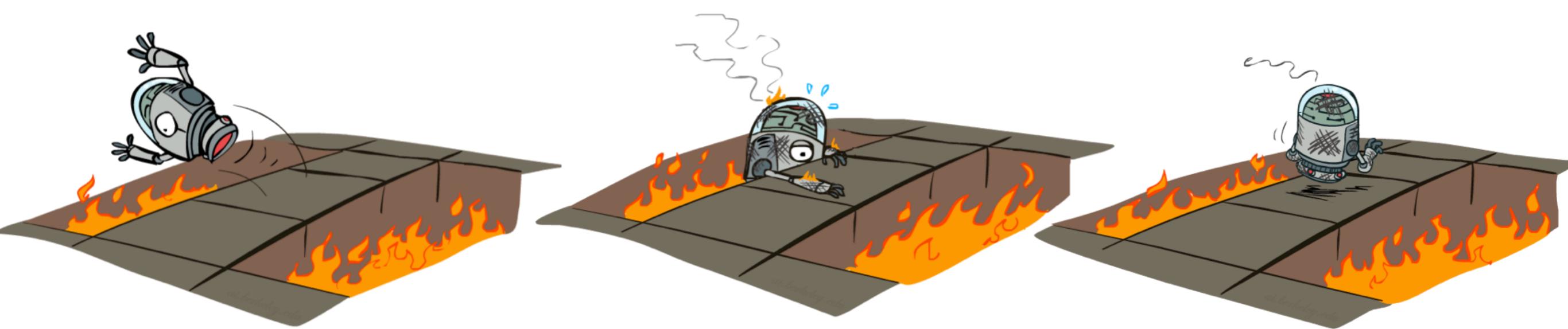
$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

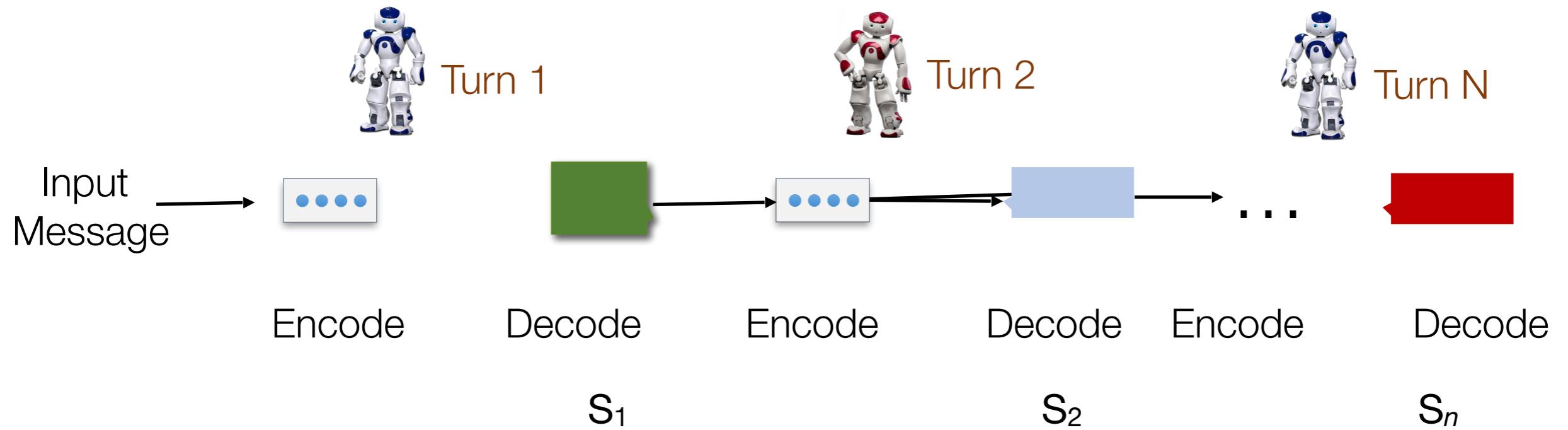




REINFORCE Algorithm  
(William, 1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

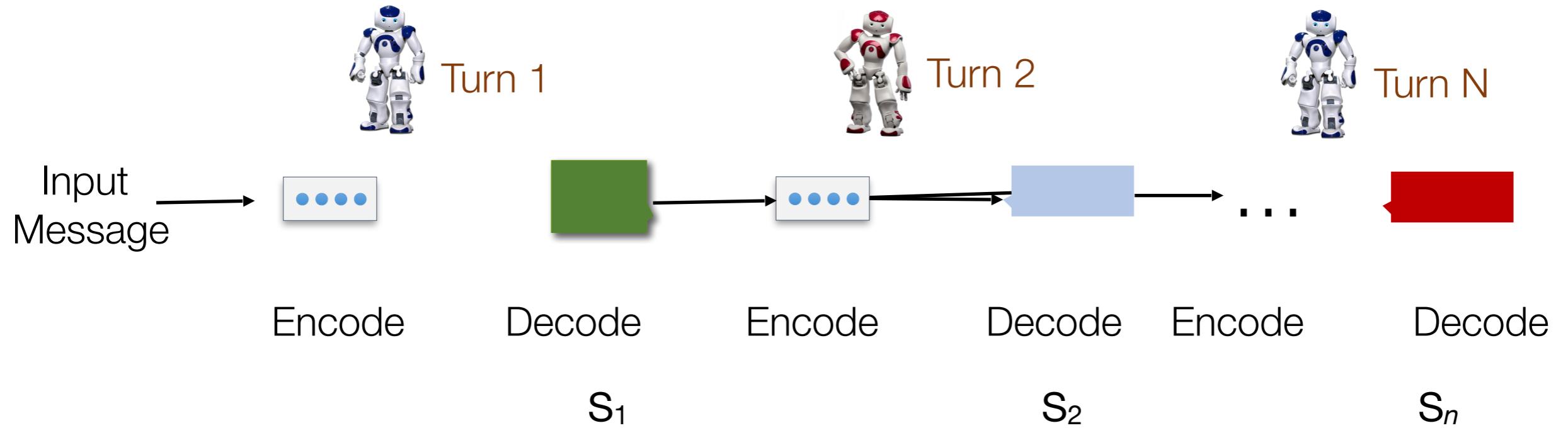




REINFORCE Algorithm  
(William, 1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

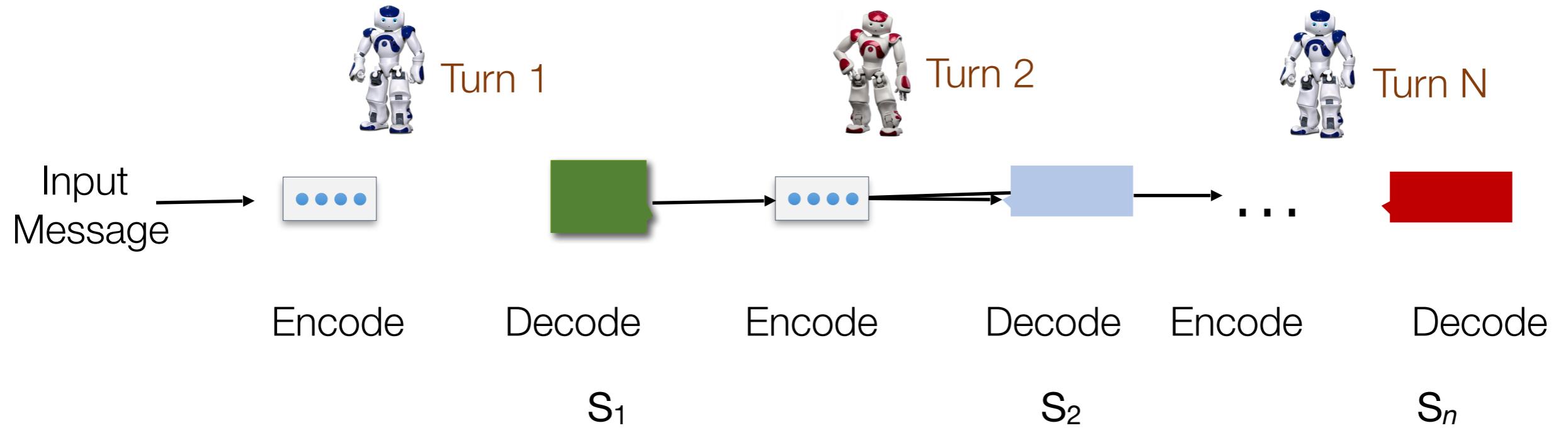
$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm  
(William, 1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

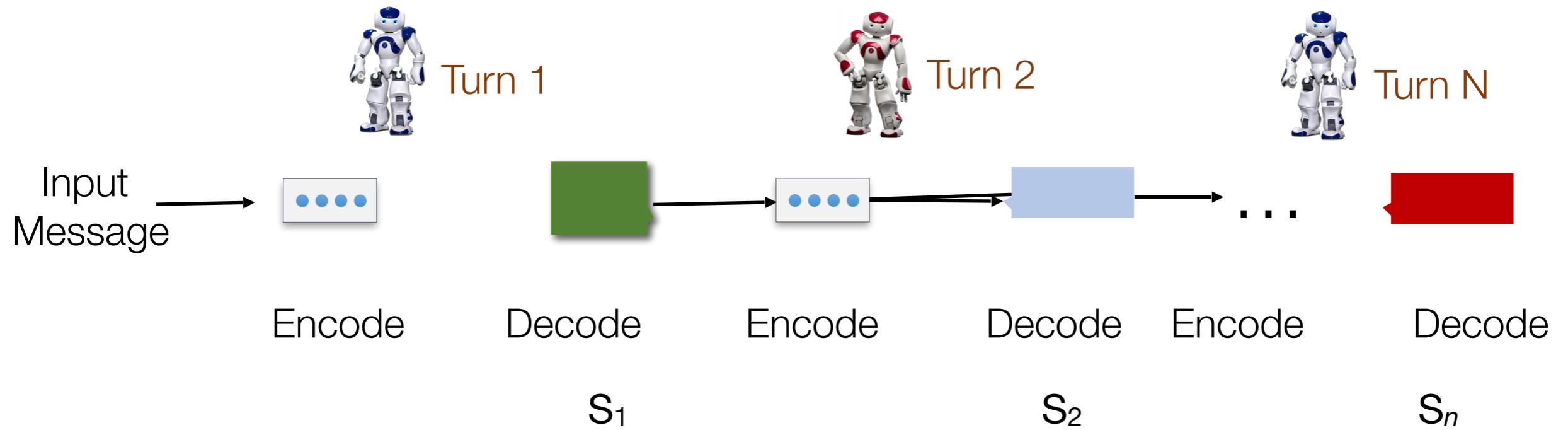
$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm  
(William, 1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$

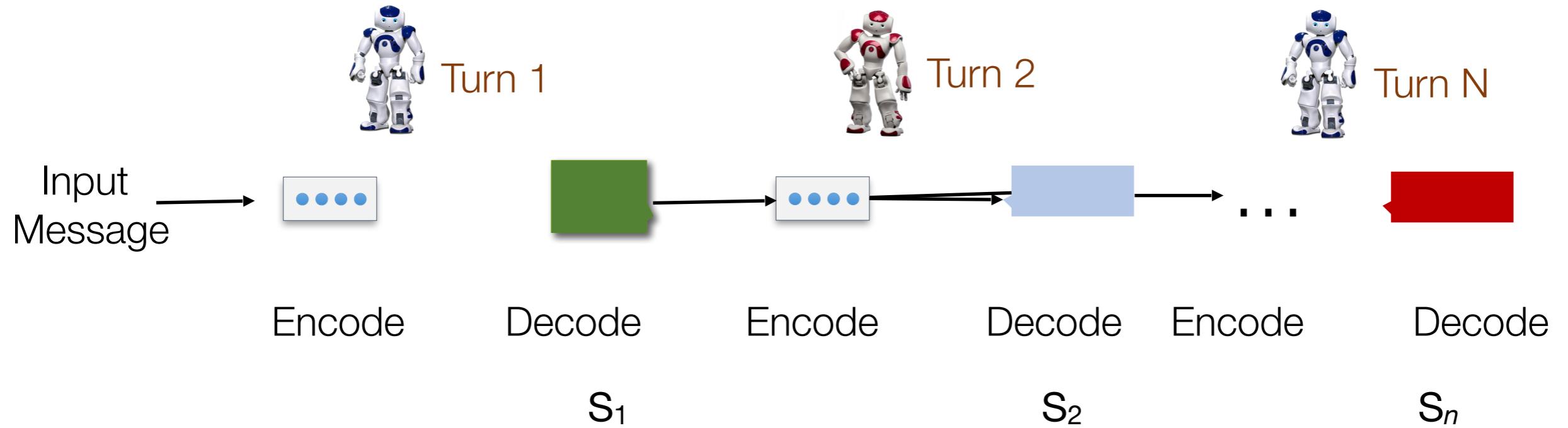


REINFORCE Algorithm  
(William, 1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$

$$\nabla J(\theta) = \nabla \log \prod_i p(s_i | s_{i-1}) R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm  
(William, 1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$

$$\nabla J(\theta) = \nabla \log \prod_i p(s_i | s_{i-1}) R(s_1, s_2, \dots, s_N)$$

What we  
want to learn

# Successful dialogue

Reward =

$$\begin{aligned} & \lambda_1 [\text{Ease of answering}] \\ & + \lambda_2 [\text{Information flow}] \\ & + \lambda_3 [\text{Semantic coherence}] \end{aligned}$$

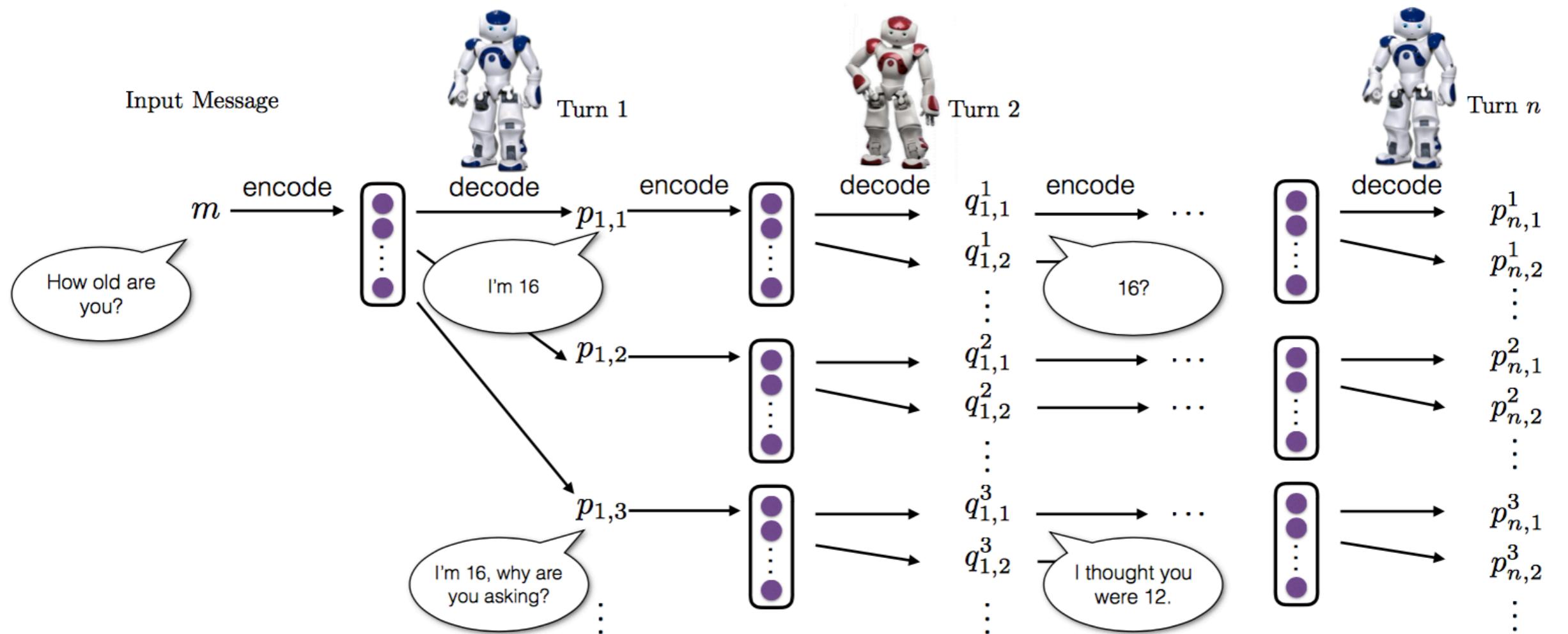


Figure 1: Dialogue simulation between the two agents.

Li et al. (2016), "Deep Reinforcement Learning for Dialogue Generation" (EMNLP)

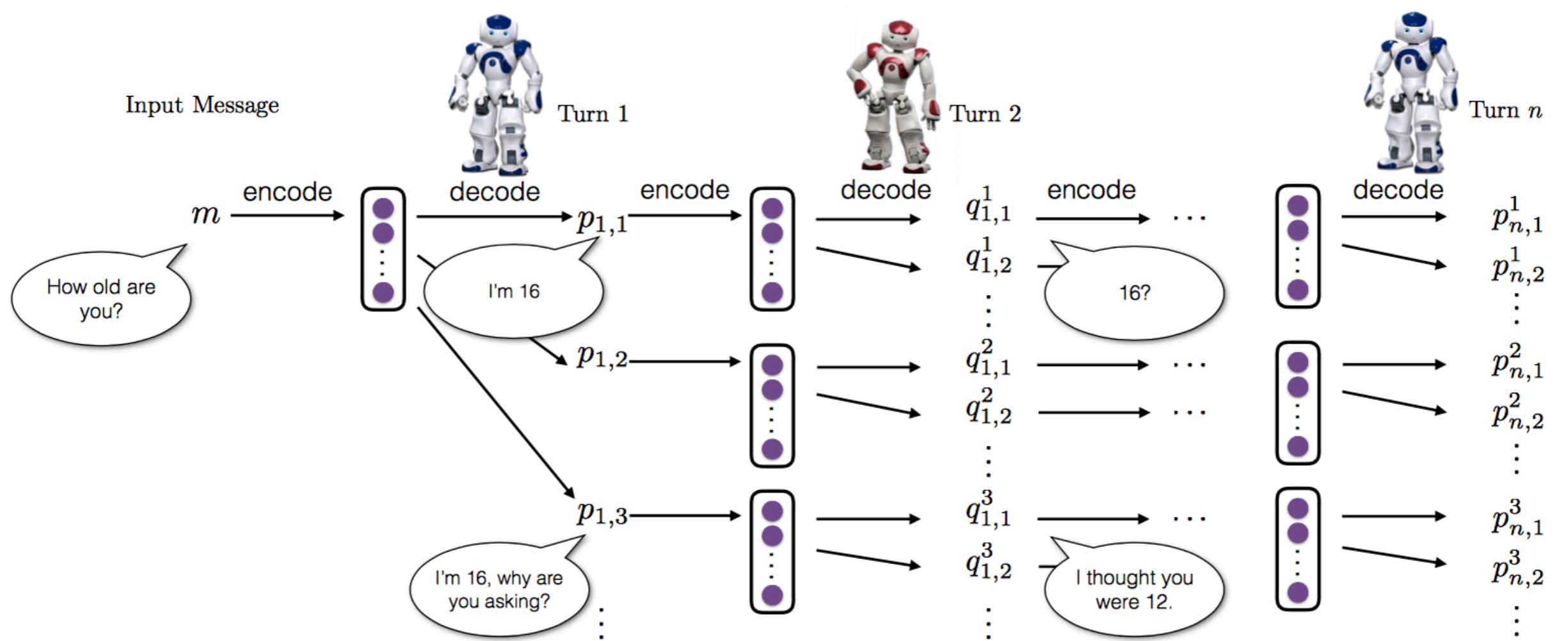


Figure 1: Dialogue simulation between the two agents.

Li et al. (2016), "Deep Reinforcement Learning for Dialogue Generation" (EMNLP)

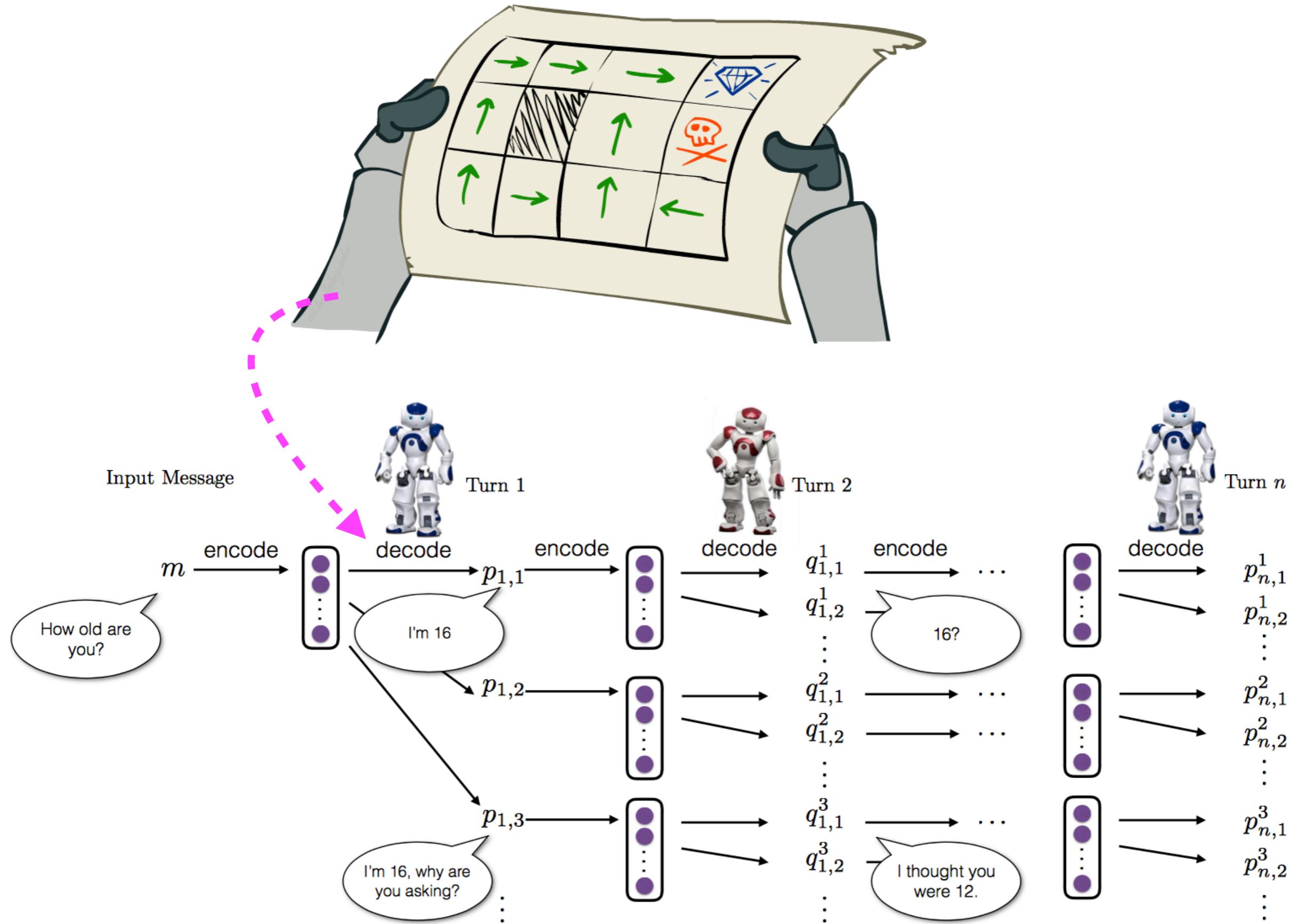


Figure 1: Dialogue simulation between the two agents.

Li et al. (2016), "Deep Reinforcement Learning for Dialogue Generation" (EMNLP)

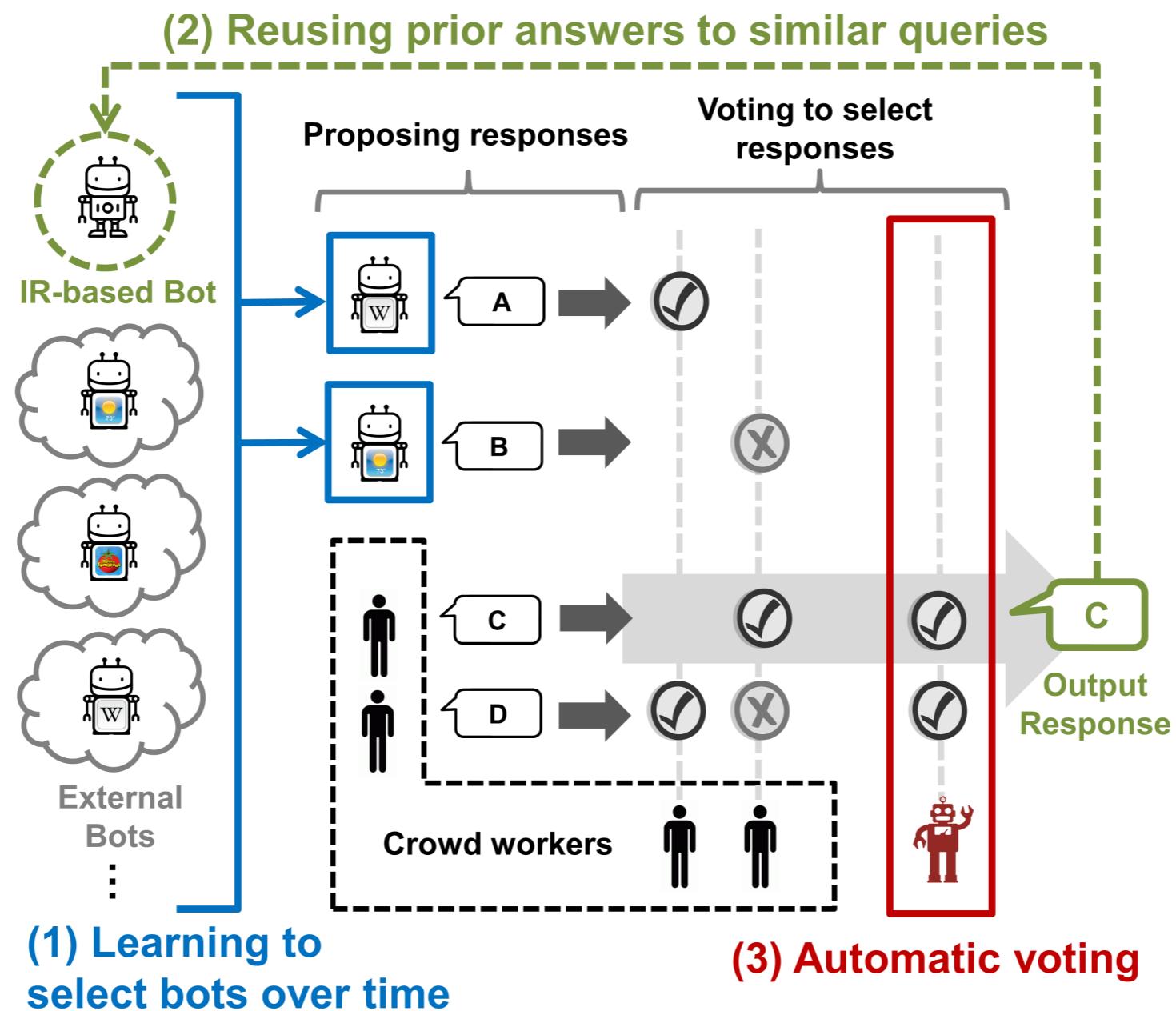
# Evaluation Results

| Model              | # of simulated turns |
|--------------------|----------------------|
| SEQ2SEQ            | 2.68                 |
| mutual information | 3.40                 |
| RL                 | 4.48                 |

# Data you could try out!

- OpenSubtitles 2018  
<http://opus.nlpl.eu/OpenSubtitles2018.php>

# Lots of HCI Cross-over Opportunities

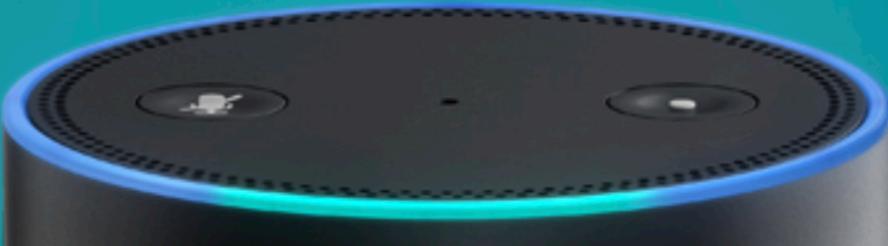


# Lots of research on dialogue agents too

## The Alexa Prize

Over \$3.5 Million to Advance Conversational Artificial Intelligence

December 2017 - November 2018



Teams of university students will develop a [socialbot](#), a new Alexa skill that converses with users on popular societal topics. Participating teams will advance the state-of-the-art in natural language understanding, dialogue and context modeling, and human-like language generation and expression.

# UNIVERSITY OF MICHIGAN - AUDREY

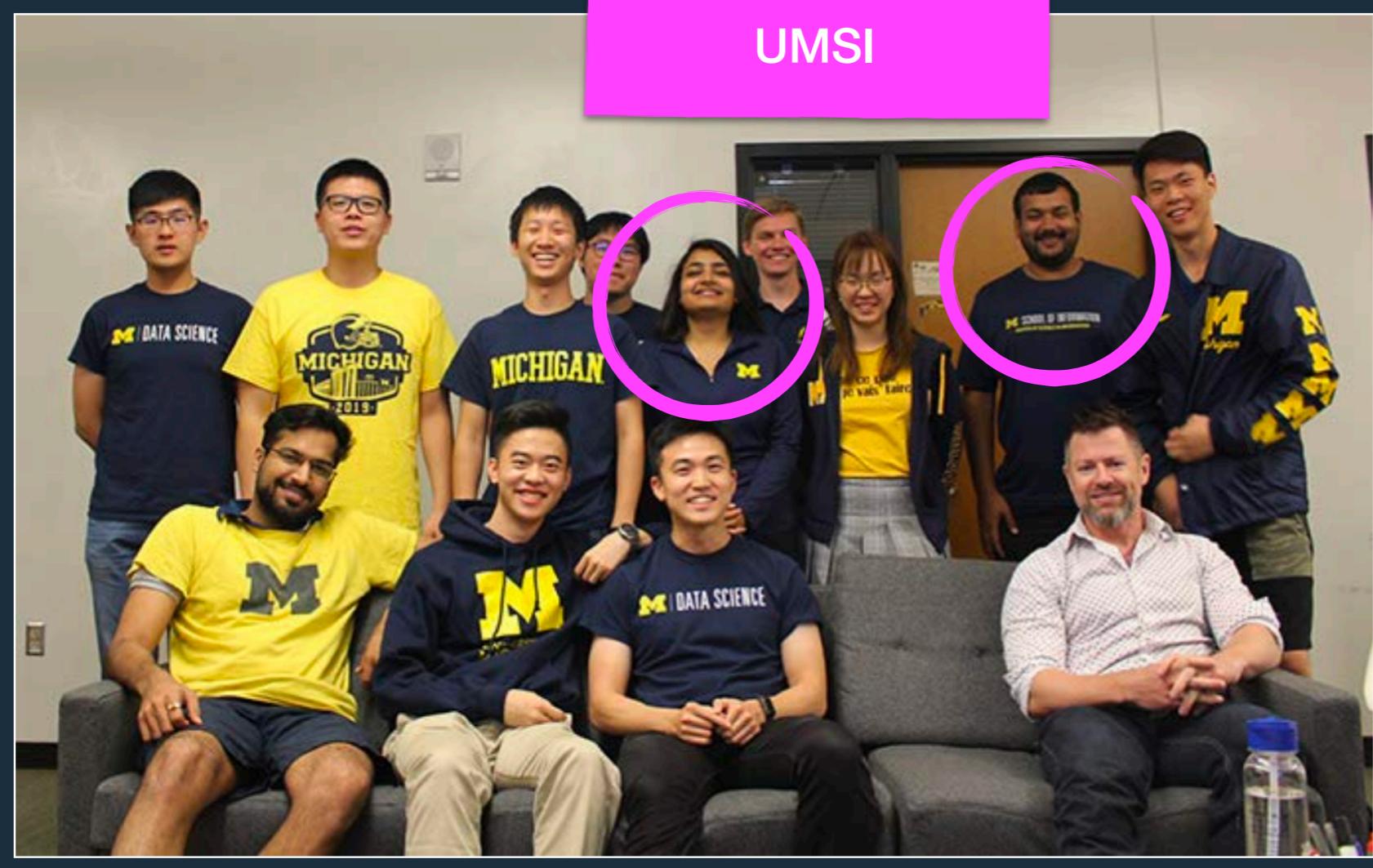


Location: Ann Arbor, MI, USA

Faculty Advisor: [David Jurgens](#)

# UNIVERSITY OF MICHIGAN - AUDREY

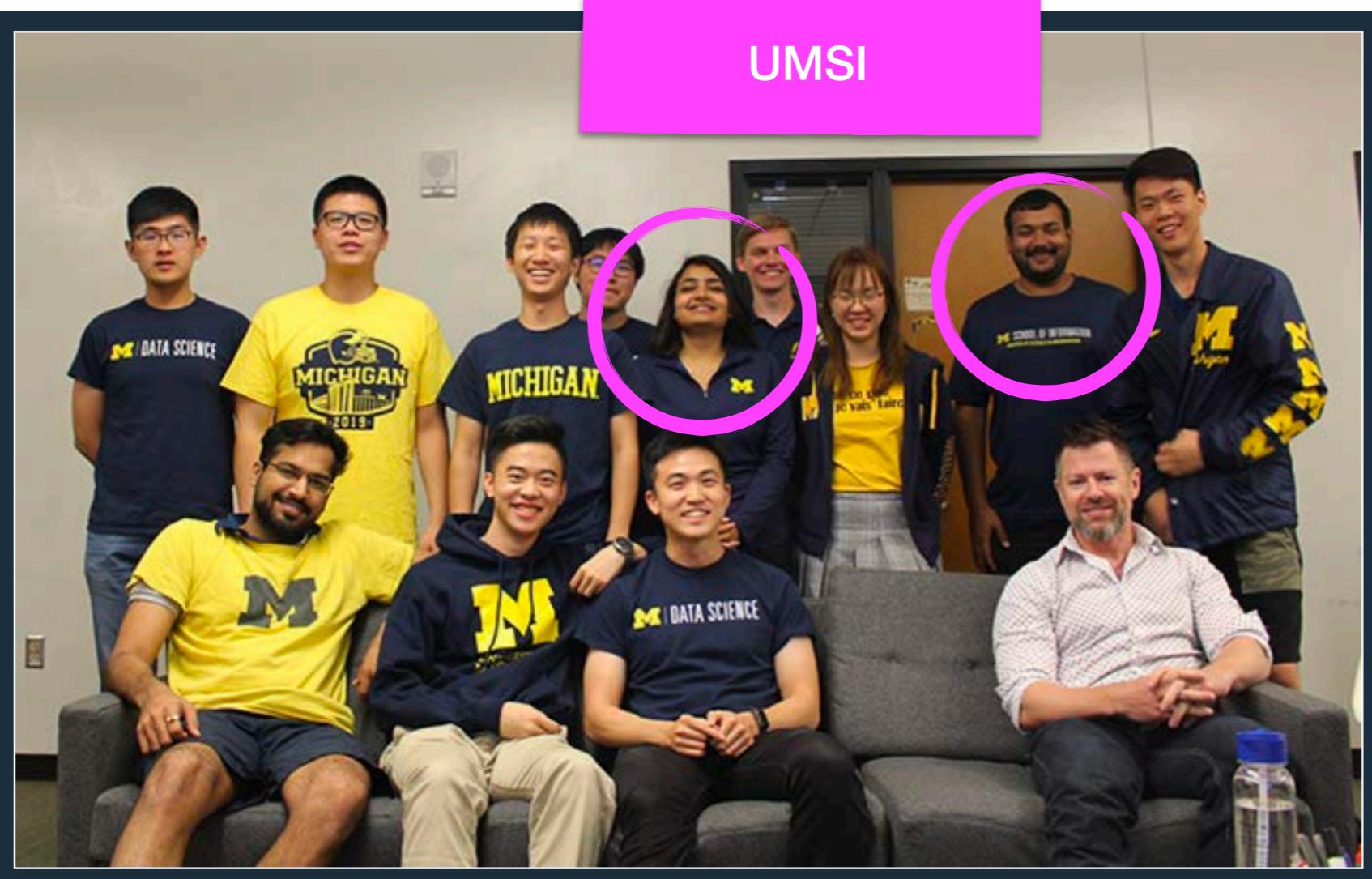
UMSI



Location: Ann Arbor, MI, USA

Faculty Advisor: [David Jurgens](#)

# UNIVERSITY OF MICHIGAN - AUDREY



UMSI

A group photo of the University of Michigan - Audrey team. There are 12 people in total, arranged in two rows. The front row has four people seated on a grey couch. The back row has eight people standing behind them. Several team members are wearing Michigan Wolverines apparel, including t-shirts with "MICHIGAN" and "DATA SCIENCE". Two individuals in the back row are circled with pink circles. A pink box labeled "UMSI" is overlaid at the top center of the image.

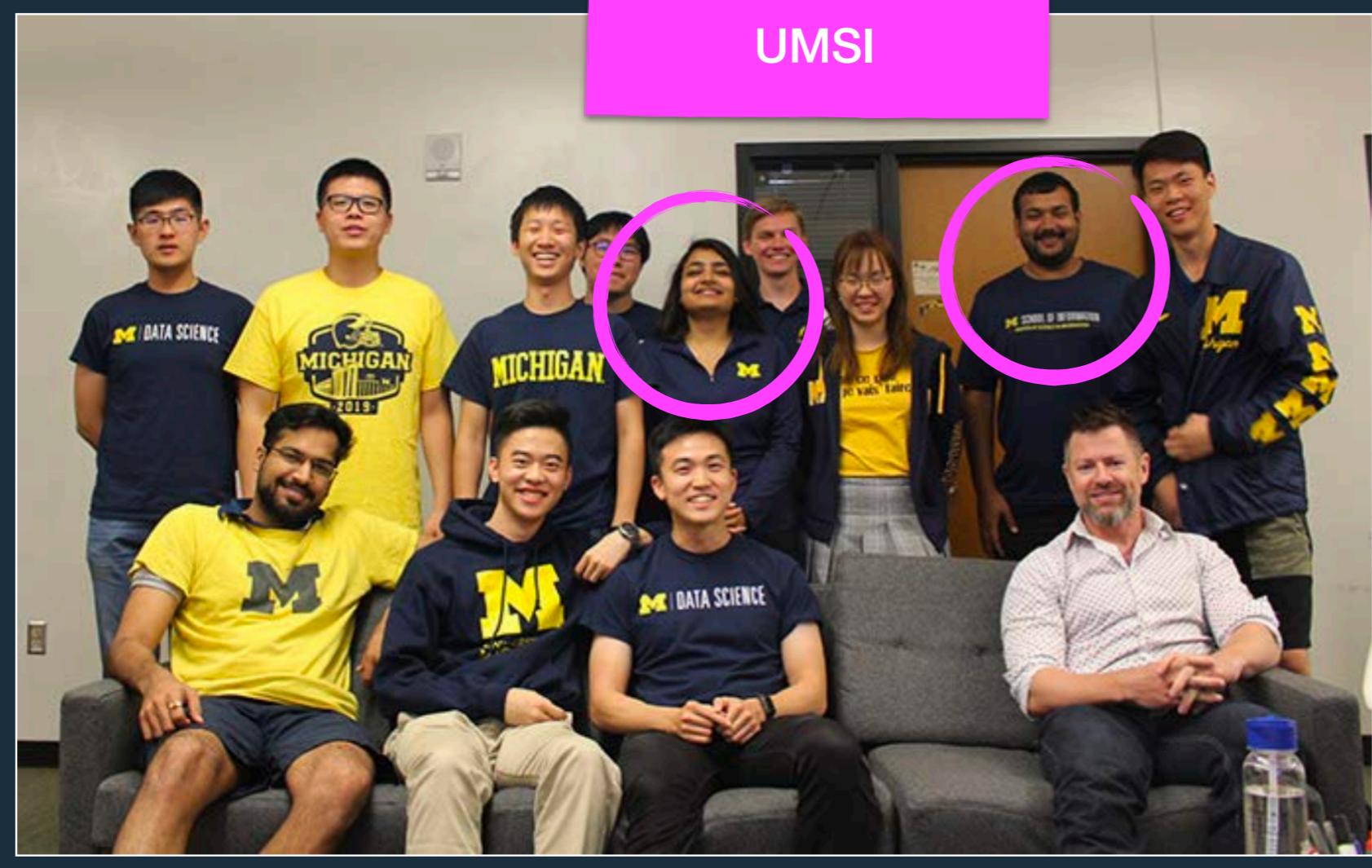
Location: Ann Arbor, MI, USA

Faculty Advisor: [David Jurgens](#)

We just got news the team made it to the semifinals!

# UNIVERSITY OF MICHIGAN - AUDREY

UMSI



Location: Ann Arbor, MI, USA

Faculty Advisor: [David Jurgens](#)

We just got news the team made it to the semifinals!

Happy to take questions on the approach



# This week's highlights

- Coreference resolution is a way of linking references (e.g., pronouns) to the name/thing they referred to
- Discourse relationships specify how discourse units (clauses, usually) relate to each other logically
  - We can use a similar setup as our parser in HW3 to find these out!
- Chatbots are hard for open-ended dialog—reinforcement learning can help make models learn how to develop good conversations

# Reminders

# Reminders

- Project Update due\* next week!

# Reminders

- Project Update due\* next week!
- HW4 due\* next week

# Reminders

- Project Update due\* next week!
- HW4 due\* next week
- Midterm in two weeks (day-of-week) poll goes out today
  - 2 hour exam that you can do over 24 hours

# Reminders

- Project Update due\* next week!
- HW4 due\* next week
- Midterm in two weeks (day-of-week) poll goes out today
  - 2 hour exam that you can do over 24 hours
- We'll do an in-class review session next week ❤️