SkillExplorer: Understanding the Behavior of Skills in Large Scale

Zhixiu Guo, Zijin Lin, Pan Li, Kai Chen

SKLOIS, Institute of Information Engineering, Chinese Academy of Sciences, China
School of Cyber Security, University of Chinese Academy of Sciences, China

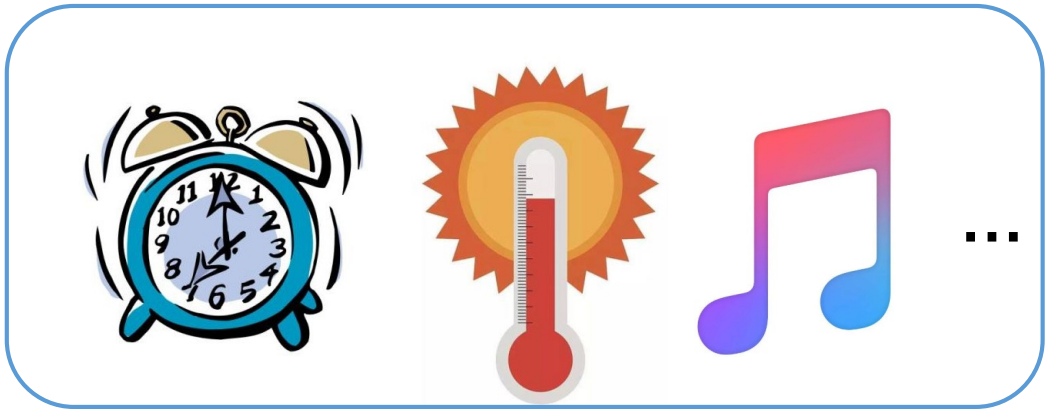
Background

Virtual personal assistant (VPA)

Background

Virtual personal assistant (VPA)

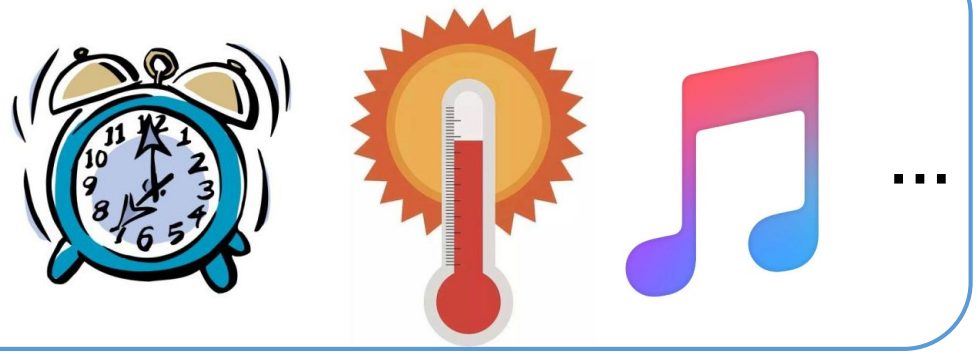
Preset



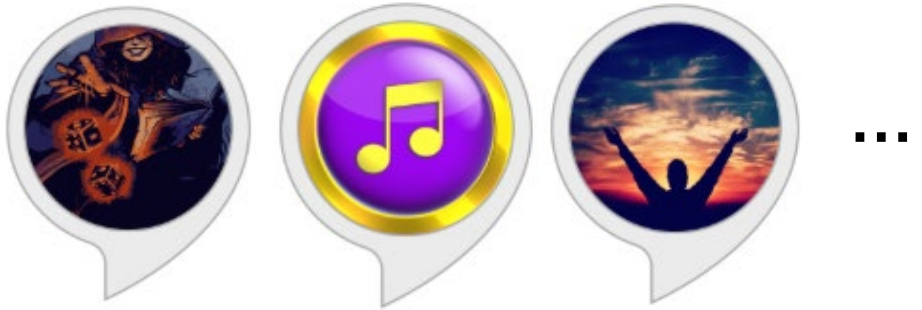
Background

Virtual personal assistant (VPA)

Preset

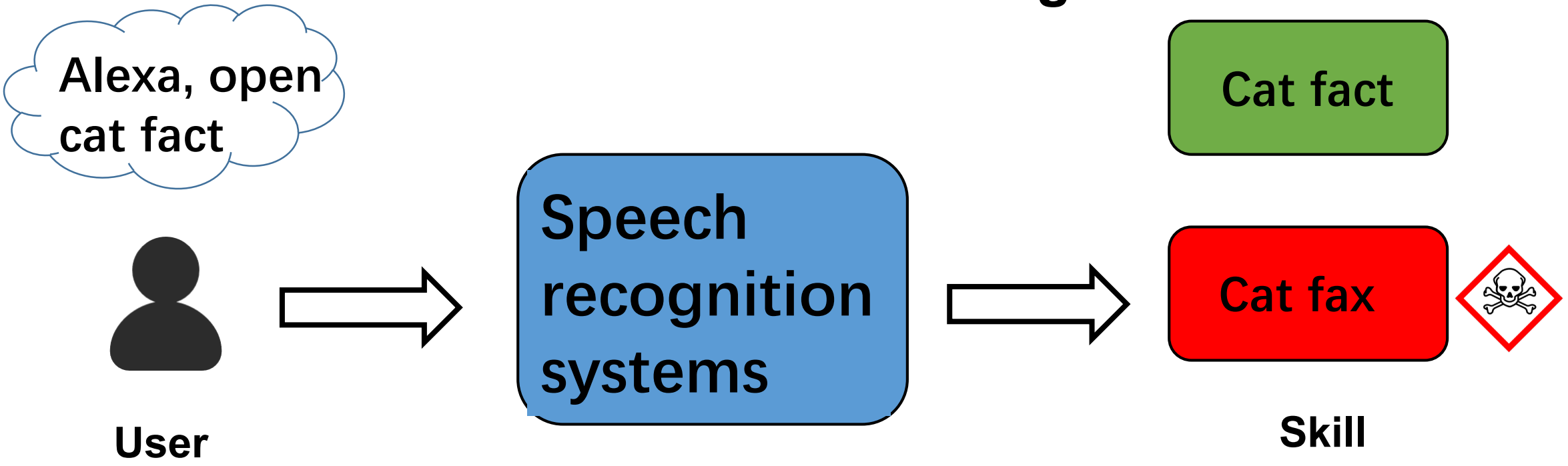


Third-party skills



Motivation

Skills are not safe enough



Skill Squatting Attacks on Amazon Alexa

Deepak Kumar, Riccardo Paccagnella, Paul Murley, Eric Hennenfent, Joshua Mason, Adam Bates, and Michael Bailey. Skill squatting attacks on amazon alexa. 27th USENIX Security Symposium, USENIX Security 2018.

Motivation

- **A skill is a voice app**
- **Recent works focus on invocation mechanism but not the content of skills**

If we can analyze the content of a skill just like analyzing a traditional app.

✓ Systematic

✓ Automatic

Motivation

- **A skill is a voice app**
- **Recent works focus on invocation mechanism but not the content of skills**

If we can analyze the content of a skill just like analyzing a traditional app.

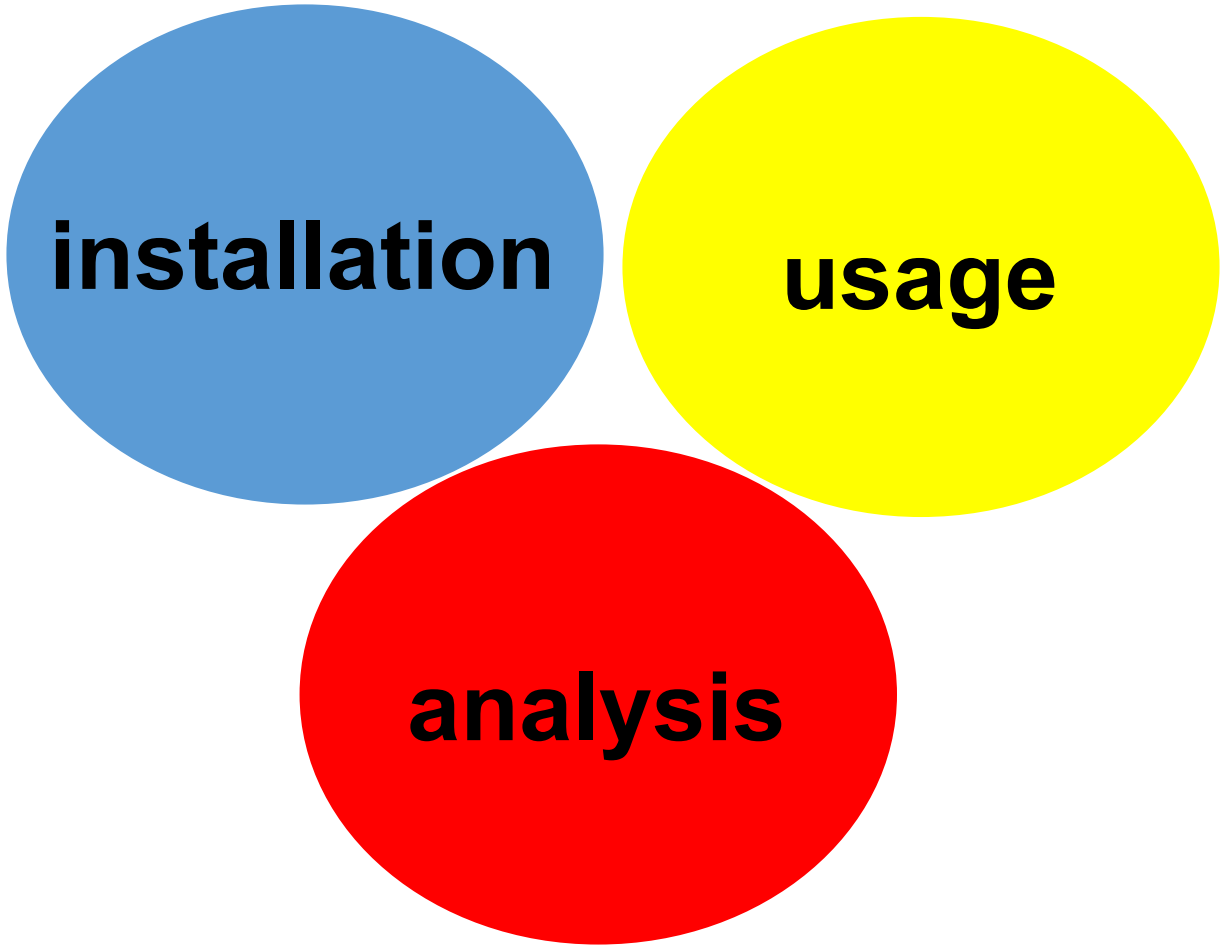
✓ Systematic

✓ Automatic

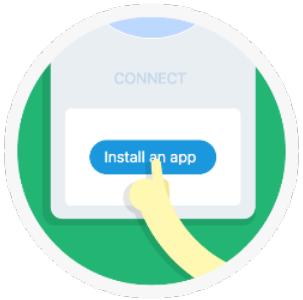
⟹ SkillExplorer

Challenges

Traditional apps and voice apps: similar but have essential differences.



Traditional Apps V.S. Voice Apps

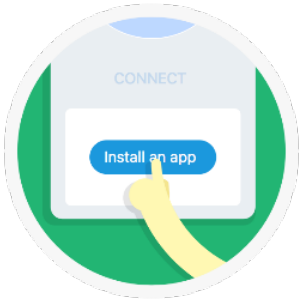


installation

usage

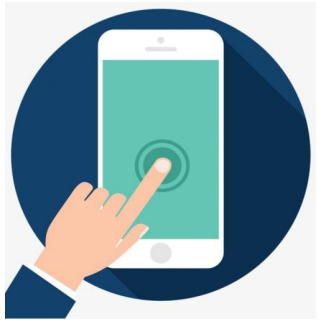
analysis

Traditional Apps V.S. Voice Apps



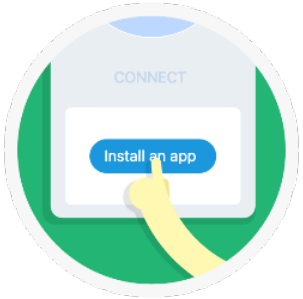
installation

usage



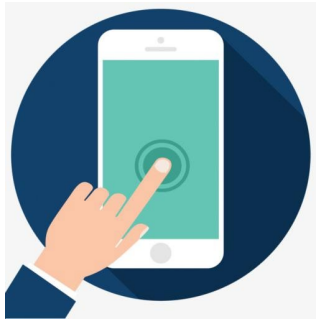
analysis

Traditional Apps V.S. Voice Apps



installation

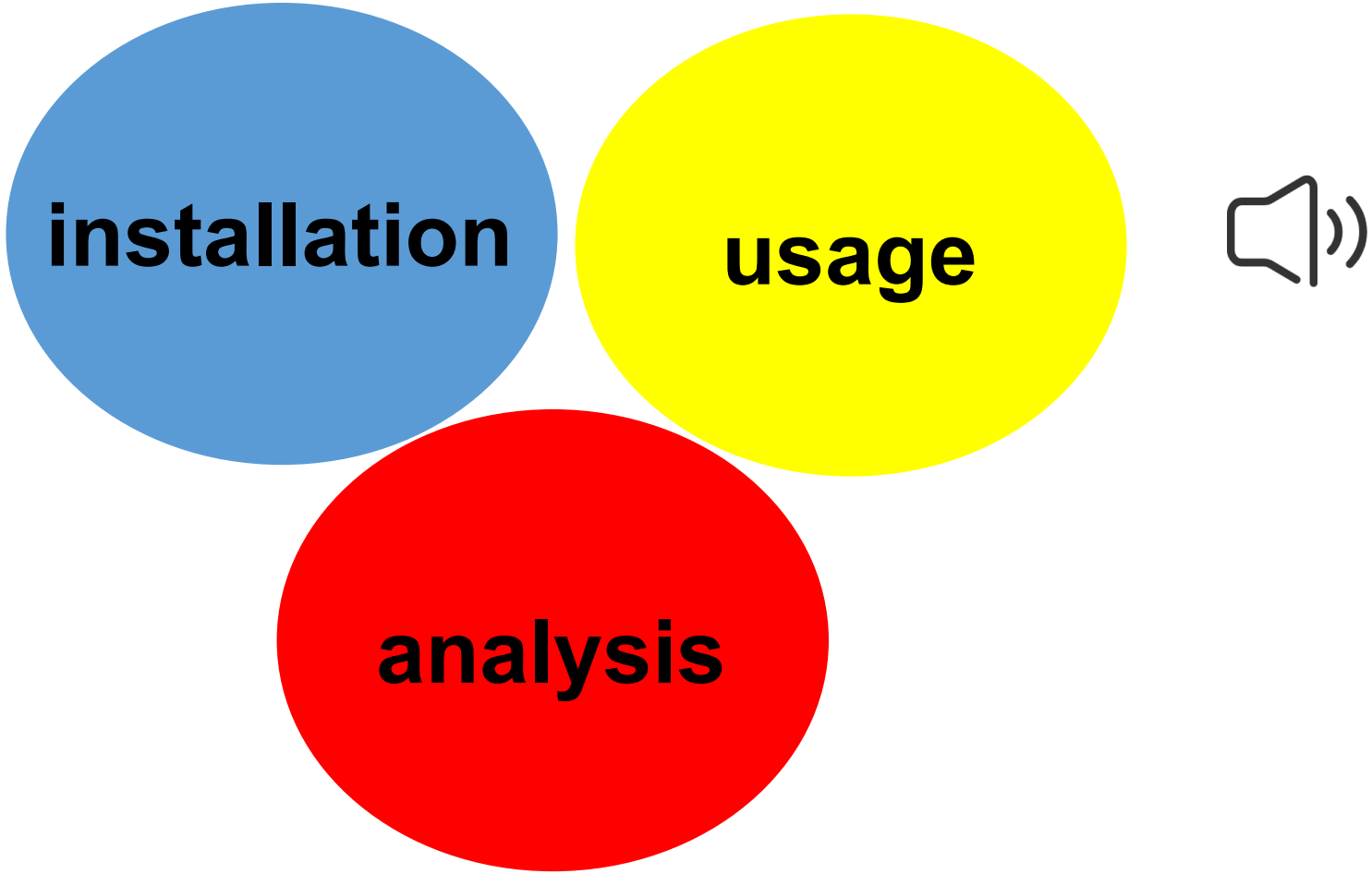
usage



analysis

Traditional Apps V.S. Voice Apps

Alexa, open/… +
invocation names



Traditional Apps V.S. Voice Apps

Alexa, open/… +
invocation names

installation

usage

analysis

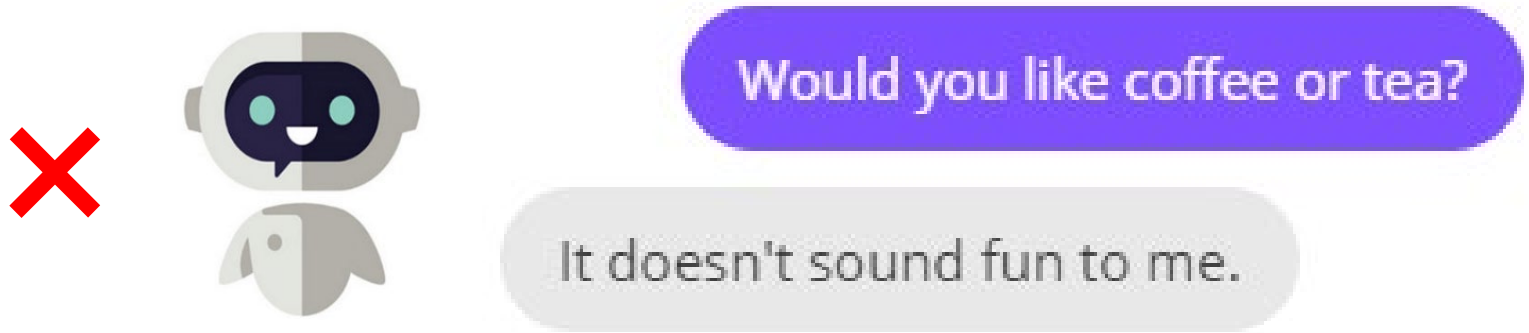
× code
× document

Challenges

- Fully black-box
- Inputs/outputs of skills are in the form of natural languages

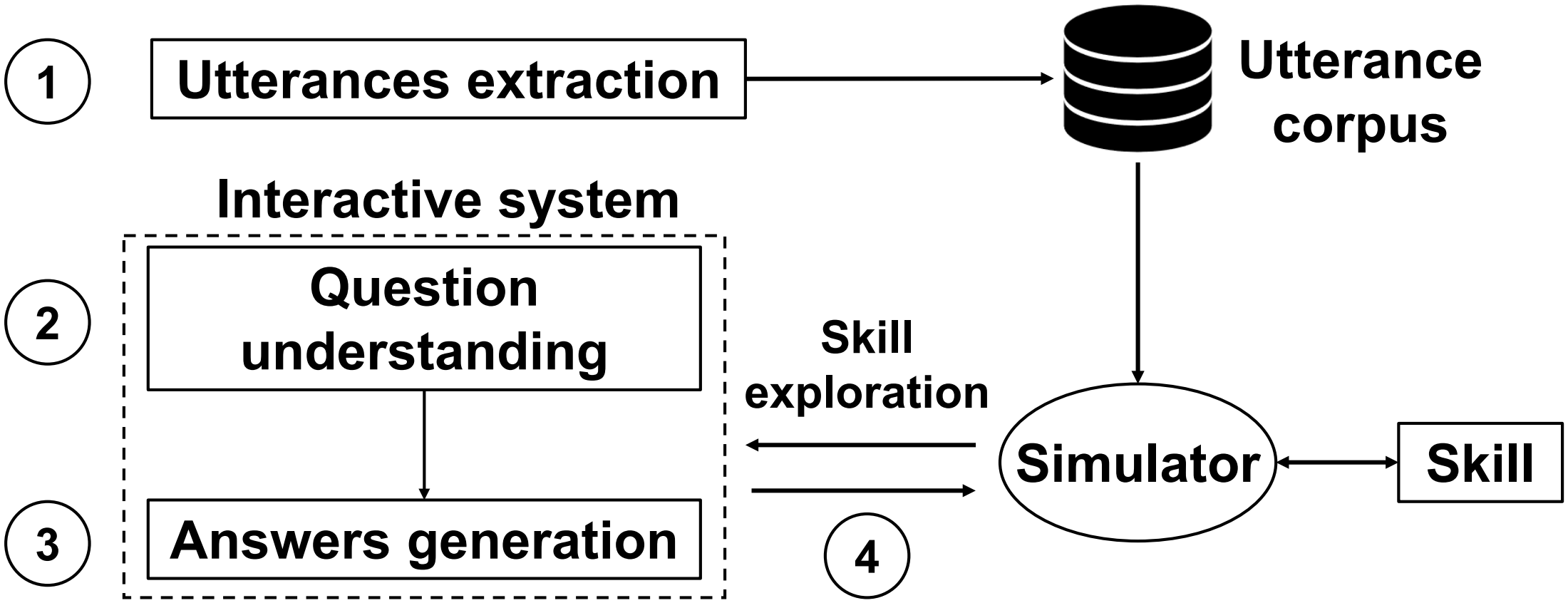
Challenges

- Fully black-box
- Inputs/outputs of skills are in the form of natural languages

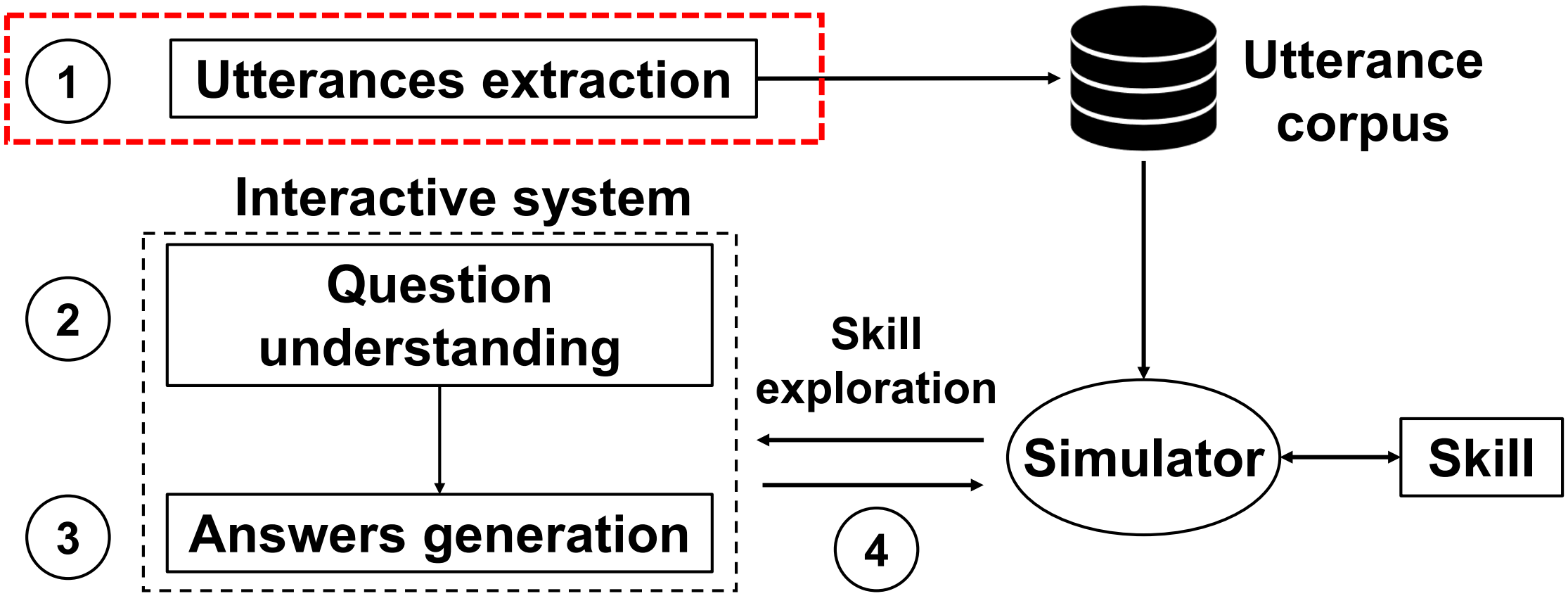


A online chatbot Mitsuku

Approach: SkillExplorer

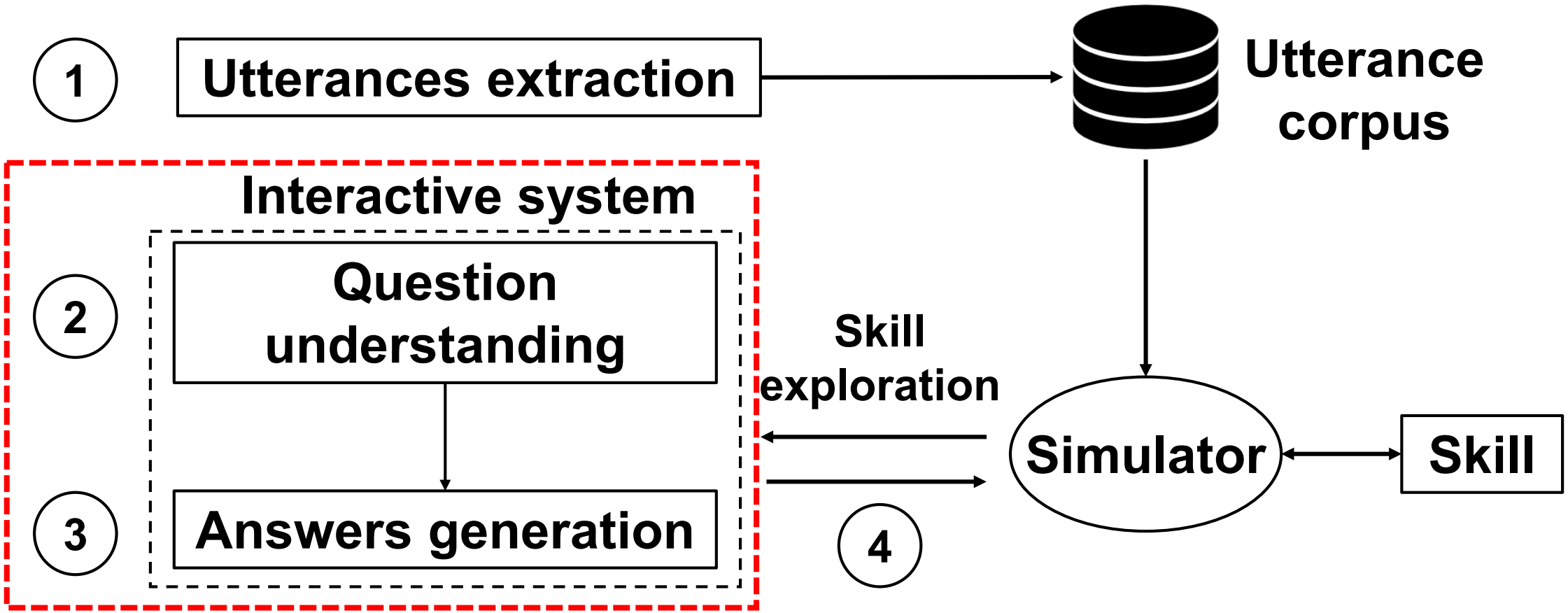


Approach: SkillExplorer

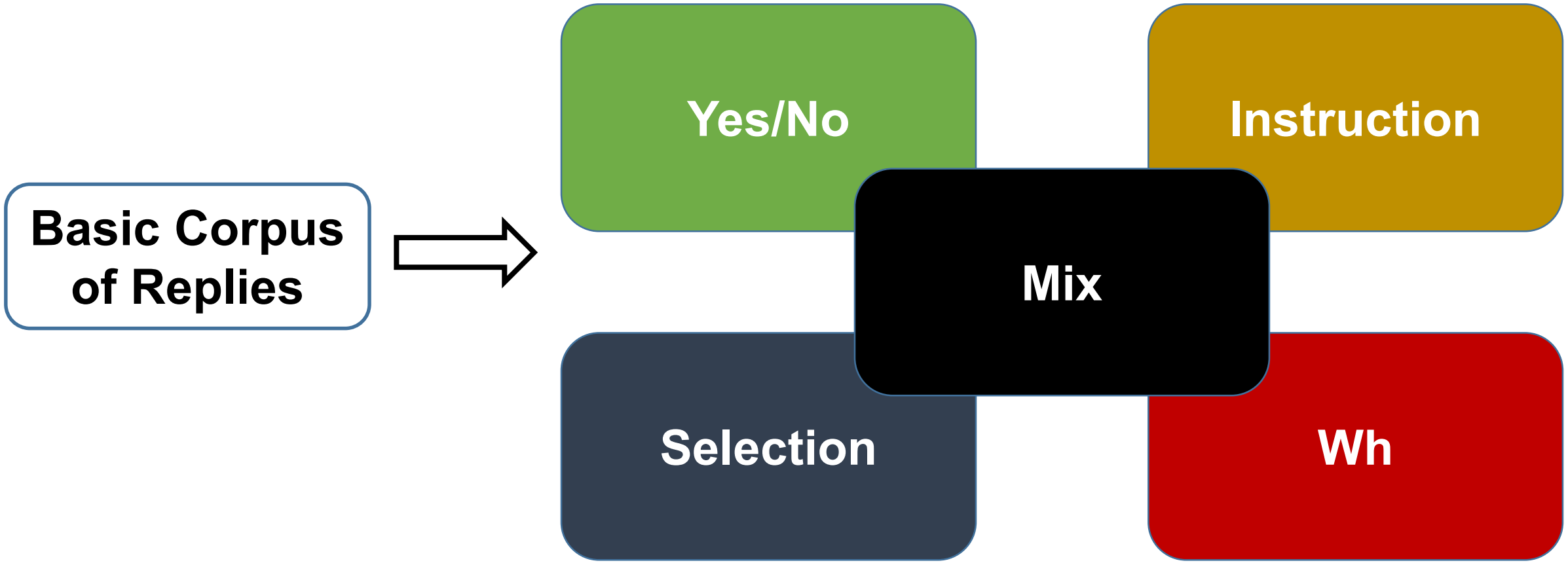


Approach: SkillExplorer

Approach: SkillExplorer



Interactive system——Five question types

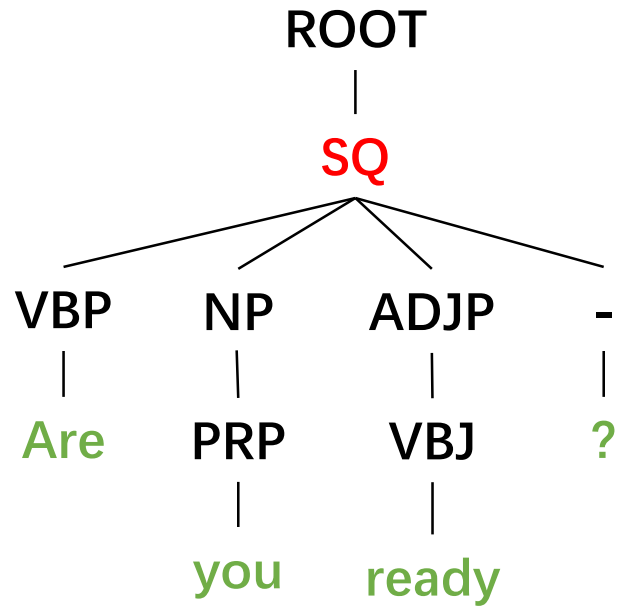


Interactive system——Five question types

Q: *Are you ready?*

A: [yes, no]

Yes/No

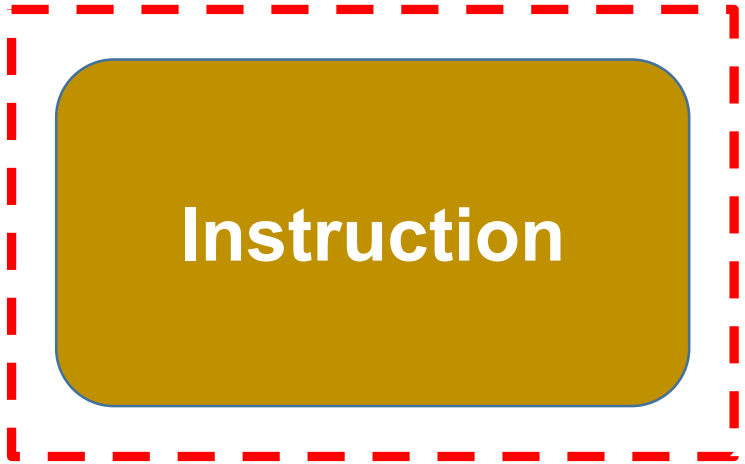


A constituency-based parse tree

Interactive system——Five question types

Q: *For any information on how to use the skill, just say: Help me.*

A: *[help me]*



ASK	SAY
ask (sb.) Wh-Q	say Wh-Q
ask sth. like/... INS	say sth. like/... INS
ask (sb.) to INS	say INS (to do sth)
ask (sb.) (about/for) INS	say INS for sth
ask that INS	say (that) INS

Rules to generate answers for Instruction questions

Interactive system——Five question types

Selection_SC (sequence numbers or sequence letters)

Q: 1: high, 2: medium, 3: low. Choose one.

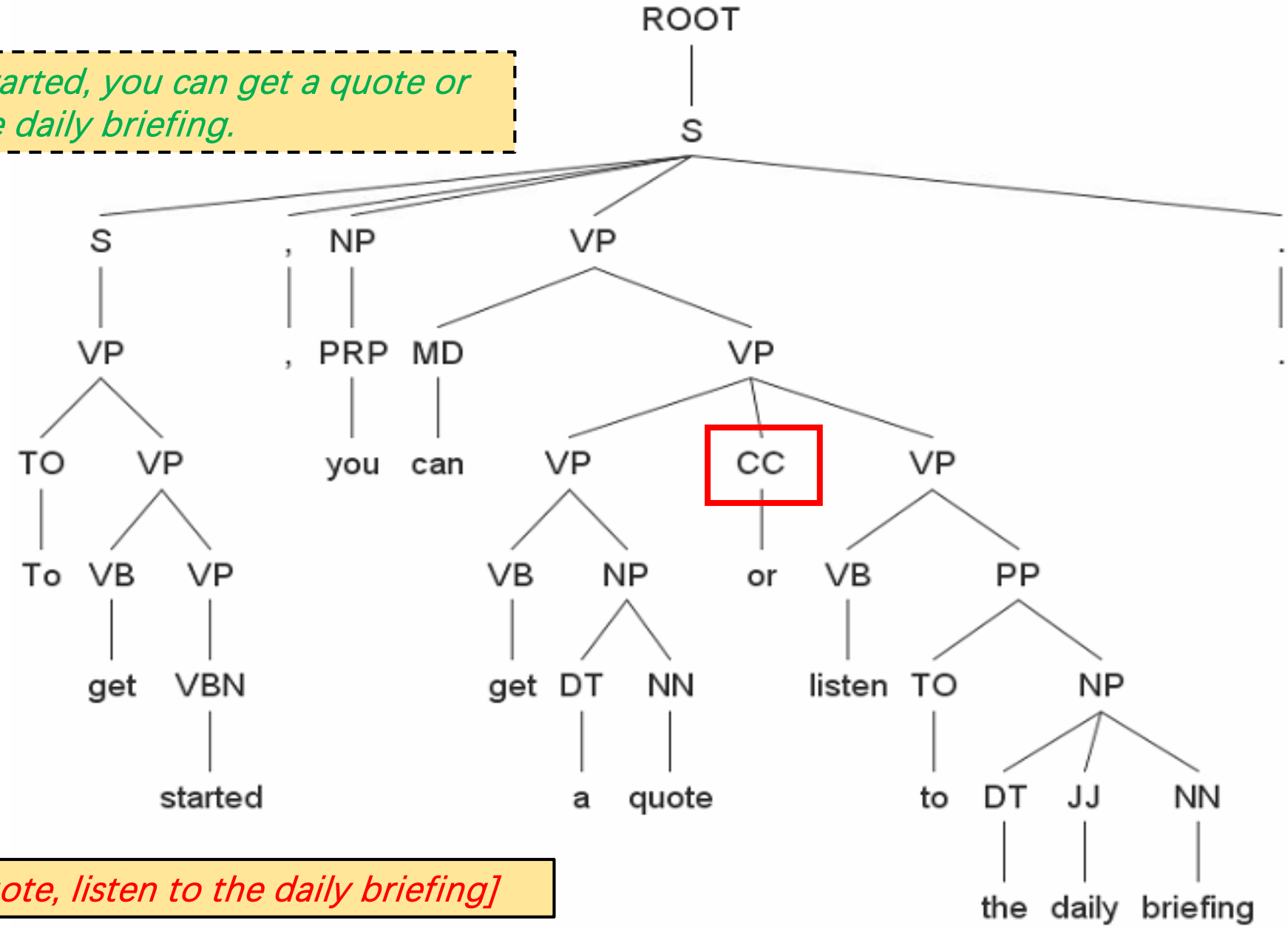
A: *[one, two, three]*

Selection_CC (coordinating conjunctions label)

Q: *To get started, you can get a quote or listen to the daily briefing.*

A: *[get a quote, listen to the daily briefing]*

Q: *To get started, you can get a quote or listen to the daily briefing.*



A: *[get a quote, listen to the daily briefing]*

A constituency-based parse tree

Interactive system——Five question types

Q: *What is your gender ?*

A: *[male, female]*

Virtual Users database

Info	Value
Full Name	James C Washington
Gender	male
Date of Birth	6/19/1980
Phone Number	716-780-4085
…	…

An example of the virtual user

Interactive system——Five question types

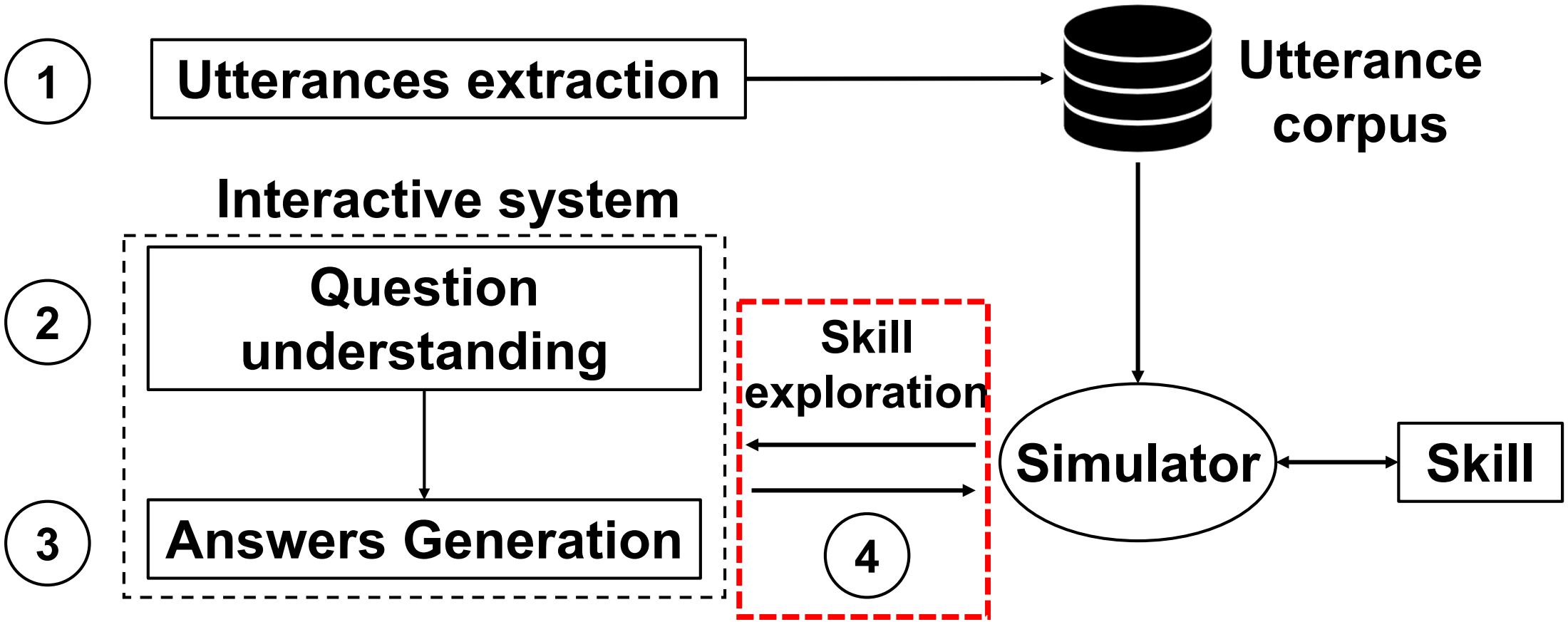
Q: *Please responds by saying lenses or glasses.*

A: [lenses, glasses]

Rule	Situation	Type
R1	∃ Y	Y
R2	∃ S_SC & ∃ I	S_SC&I
R3	(I&S_CC) in Q *	I&SC_CC
R4	∃ I	I
R5	∃ S	S

Rules to generate answers for Mix questions

Approach: SkillExplorer



Approach: SkillExplorer

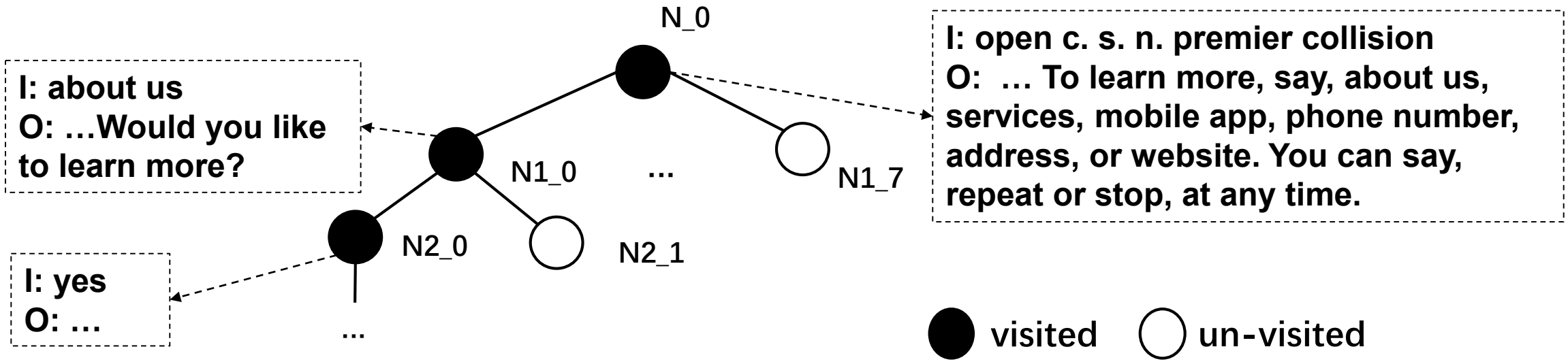
An interactive tree (i-tree for short) is used to record the exploration

- Each node represents a single interaction (include an input and a output)
- Different answers produce different branches
- The node will be marked as visited if it is explored
- Re-start from the beginning to the unvisited nodes

Approach: SkillExplorer

An interactive tree (i-tree for short) is used to record the exploration

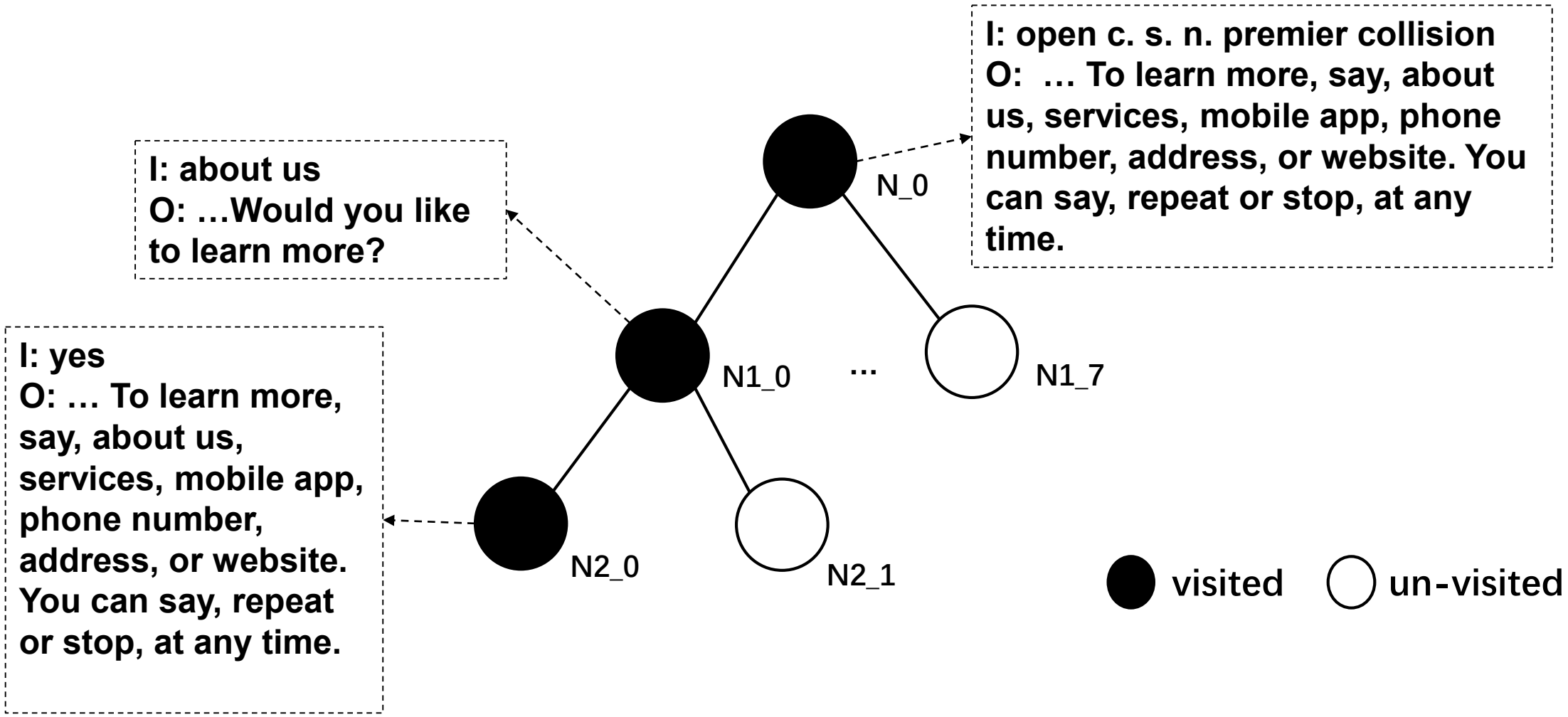
- Each node represents a single interaction (include an input and a output)
- Different answers produce different branches
- The node will be marked as visited if it is explored
- Re-start from the beginning to the unvisited nodes



An example of i-tree

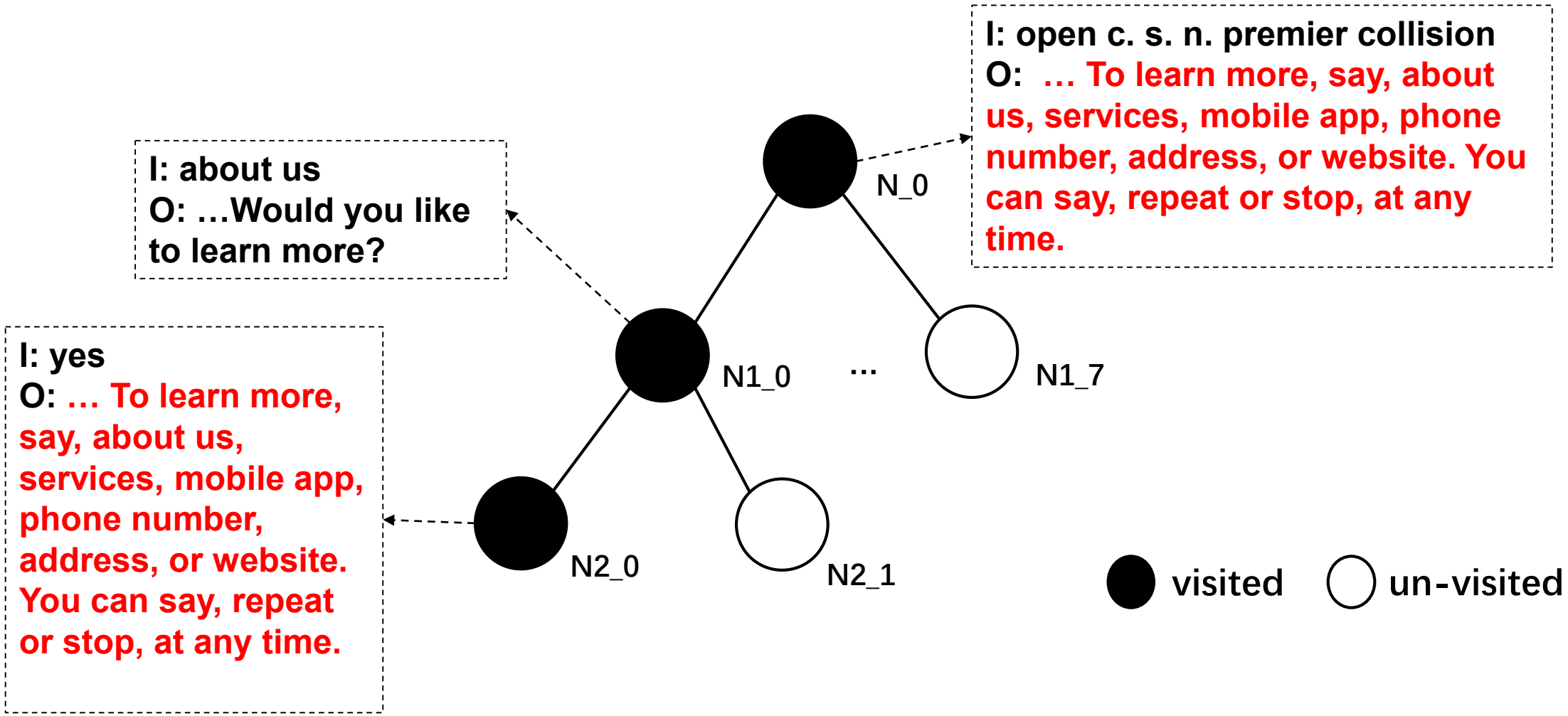
Speed up mechanism

- ignore the same questions (i.e. outputs) in different nodes
- not wait for every output reading



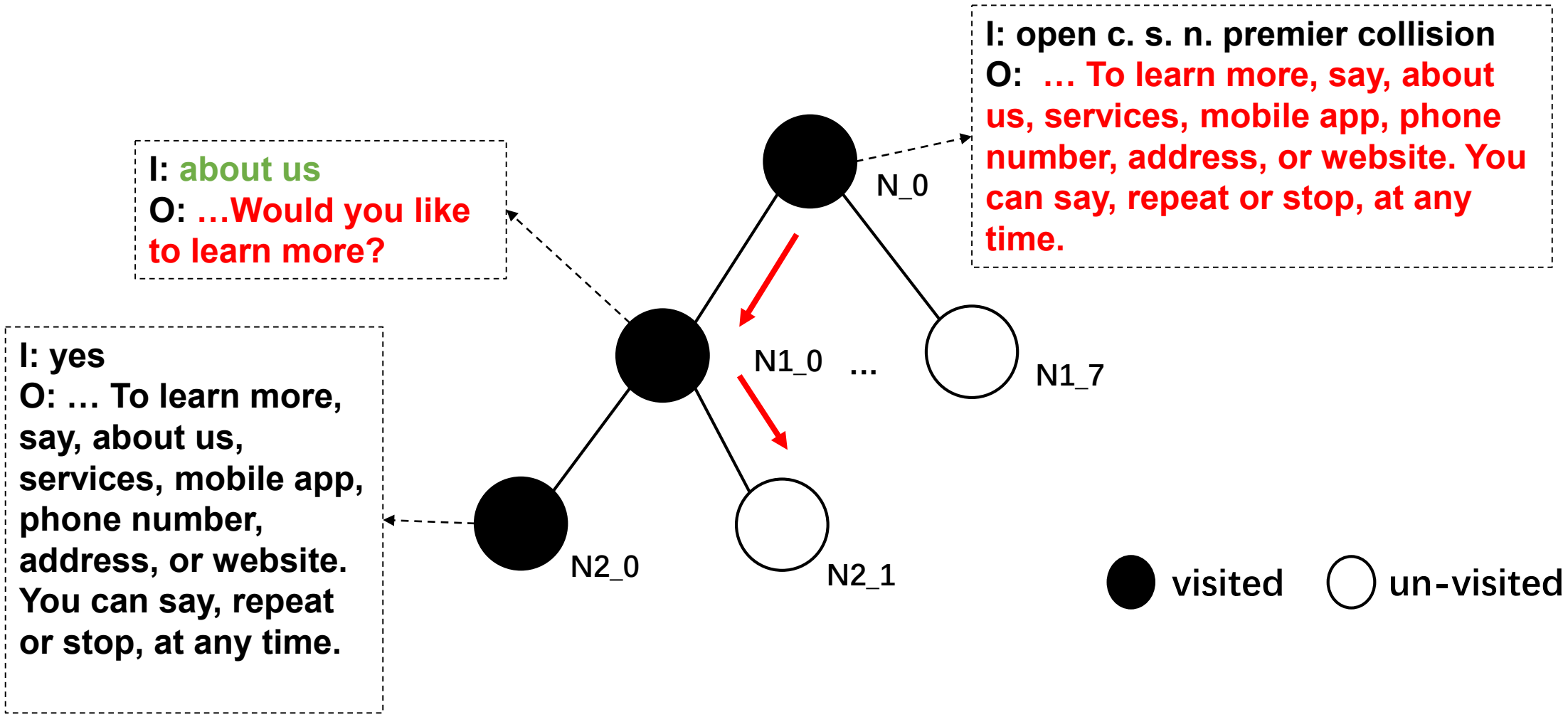
Speed up mechanism

- ignore the same questions (i.e. outputs) in different nodes
- not wait for every output reading



Speed up mechanism

- ignore the same questions (i.e. outputs) in different nodes
- not wait for every output reading



Evaluation

- **28, 904** skills from Amazon and **1,897** actions from Google through simulators
- Paths coverage rate: **90%**
- Error rate of answer generation:

Yes/No	Instruction	Selection	Wh	Mix
0%	8%	8%	5%	9%

- More than **5,200** hours spent. Each skill costs about 627 seconds on average
- The speed up mechanism saves **29.2 %** time

Findings

Landscape

- **Skills & authors**
  - 68,066 skills from Amazon market with 12,376 different developer names
  - Some developers own more than 1,000 skills

- **Invocation names**
  - 9,799 skills' invocation names do not meet Amazon's requirements (e.g. using place/people names)
  - 2591 invocation names are not unique

Findings

Developer Specifications for request private information

a) using specific APIs (e.g., Alexa customer profile API) to request permissions

b) in the privacy policy of the skills

Findings

Developer Specifications for request private information

a) using specific APIs (e.g., Alexa customer profile API) to request permissions

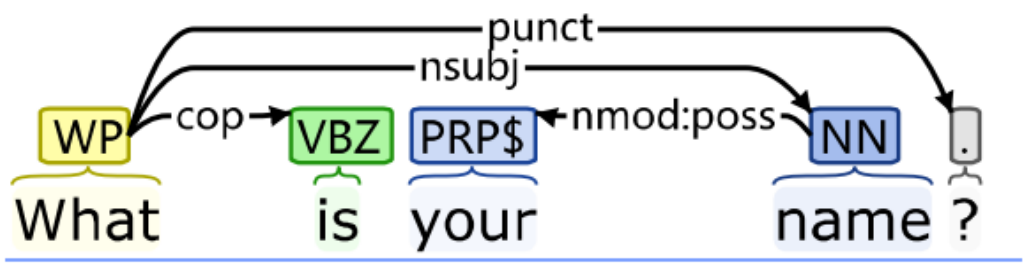
b) in the privacy policy of the skills

Method

interactive records

Name, phone number, …

- **Privacy Policy (PP)**
- **Permissions**



An example of dependency-based parse structure

Collect verb (e.g. collect, use, …)

General term (e.g. personal data …)

Subsumptive relationships (e.g. such as, …)

key components of the general declaration in PP

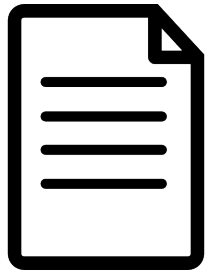
Findings

Developer Specifications for request private information

a) using specific APIs (e.g., Alexa customer profile API) to request permissions

b) in the privacy policy of the skills

Method



interactive records

Name, phone number, …

- **Privacy Policy (PP)**
- **Permissions**

Results

- **1,141** skills conflict with the developer specifications

Findings

- **68 skills** have problems when users say “stop”
  - 32 skills change the default “stop” commands (e.g. I’ve done)
  - 29 skills ignore the stop command
  - 7 skills seem more strange

Conclusion

- Develop a systematic method to explore skills
  - a suite of grammar-based approaches
- Conduct a large scale of testing in the skill market
  - about 30,000 skills from markets
- Find a good number of skills that don't follow the development rules
  - 1,141 skills request private information and 9,799 skills' invocation names
- Find some suspicious skills
  - 68 skills have problems when receiving Stop command

Thanks for listening!

Q&A

guozhixiu@iie.ac.cn