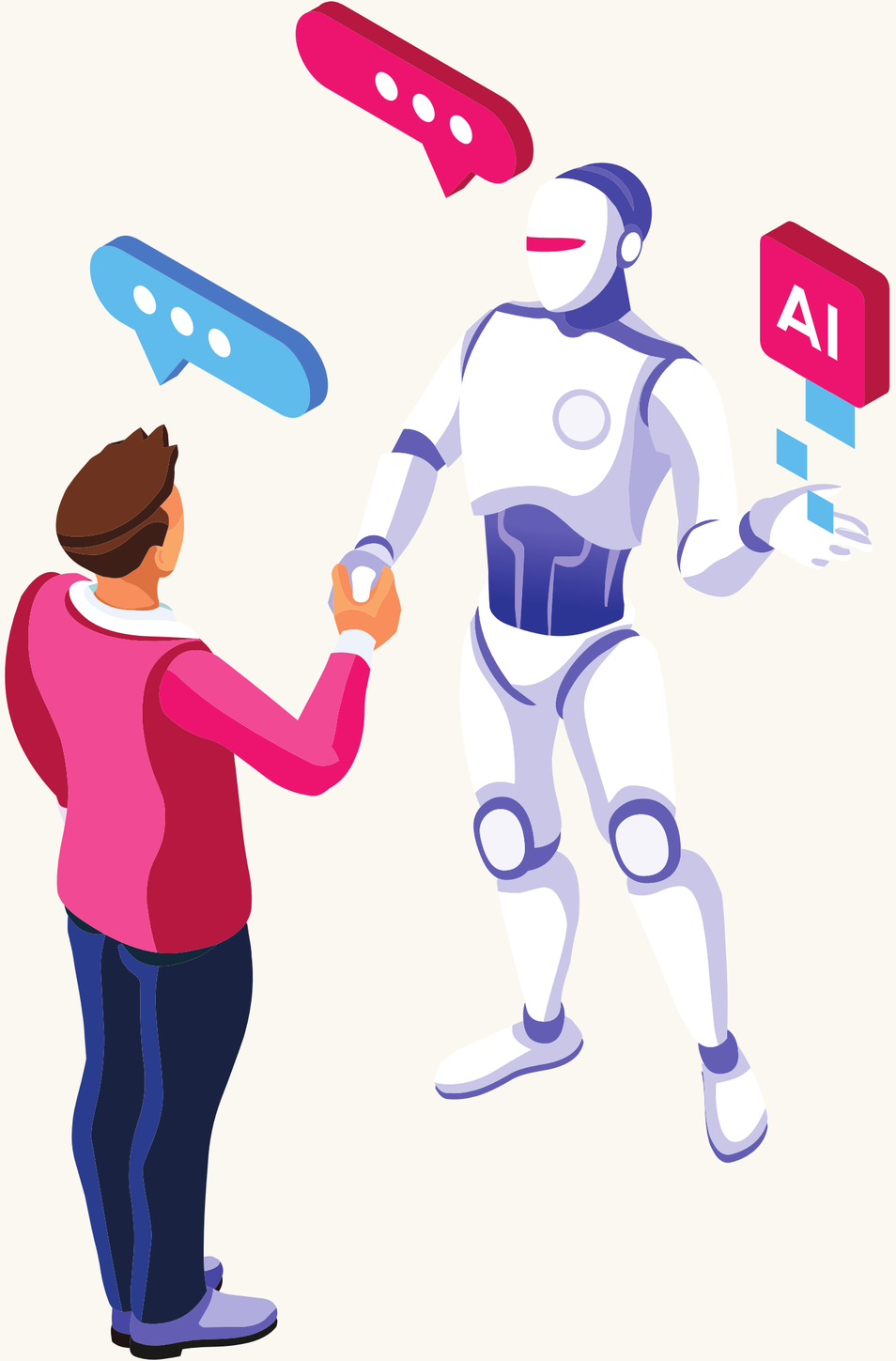


TASK ORITENTED DIALOGUE

Symbolic Dialouge for General Domain State Tracking

ACHIEVING A PRE-DEFINED TASK THROUGH A DIALOG. (PAPER WITH CODE)

Table Of Contents

Meet The Group

Nguyễn Sơn Tùng

Nguyễn Mạnh Tường

Lê Xuân Tùng

ACKNOWLEDGMENT

- We deeply appreciate support from Gradient Technologies and in-depth technical discussion with the board, especially Mr. Nguyen Van Tien

- Thank you our mentor Nguyen Quoc Trung for his dedicated guidance for the project

Conversational Chatbots

Open Domain

TOD

Multi Domain

Single Domain

Natural Language

Flexible
Ambiguous

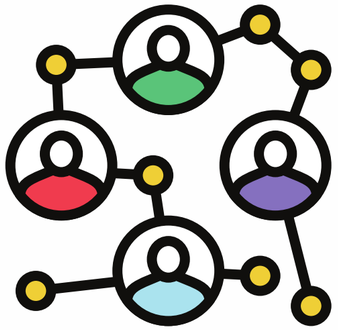
Programing Language

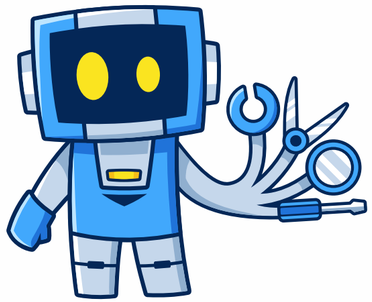
Concrete
Uninformative

SYMBOLIC

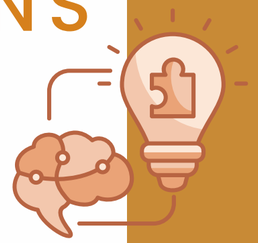
Flexible and concrete

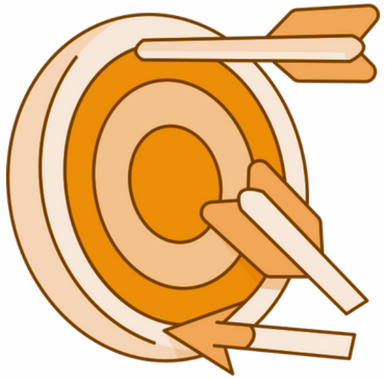
Modern LLMs





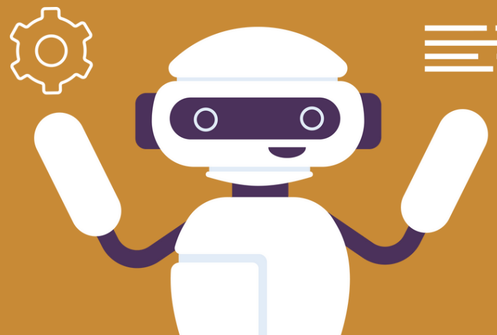
AGENTS
AND
PLUGGINS





Task Oriented Dialogue (TOD) and Symbolic

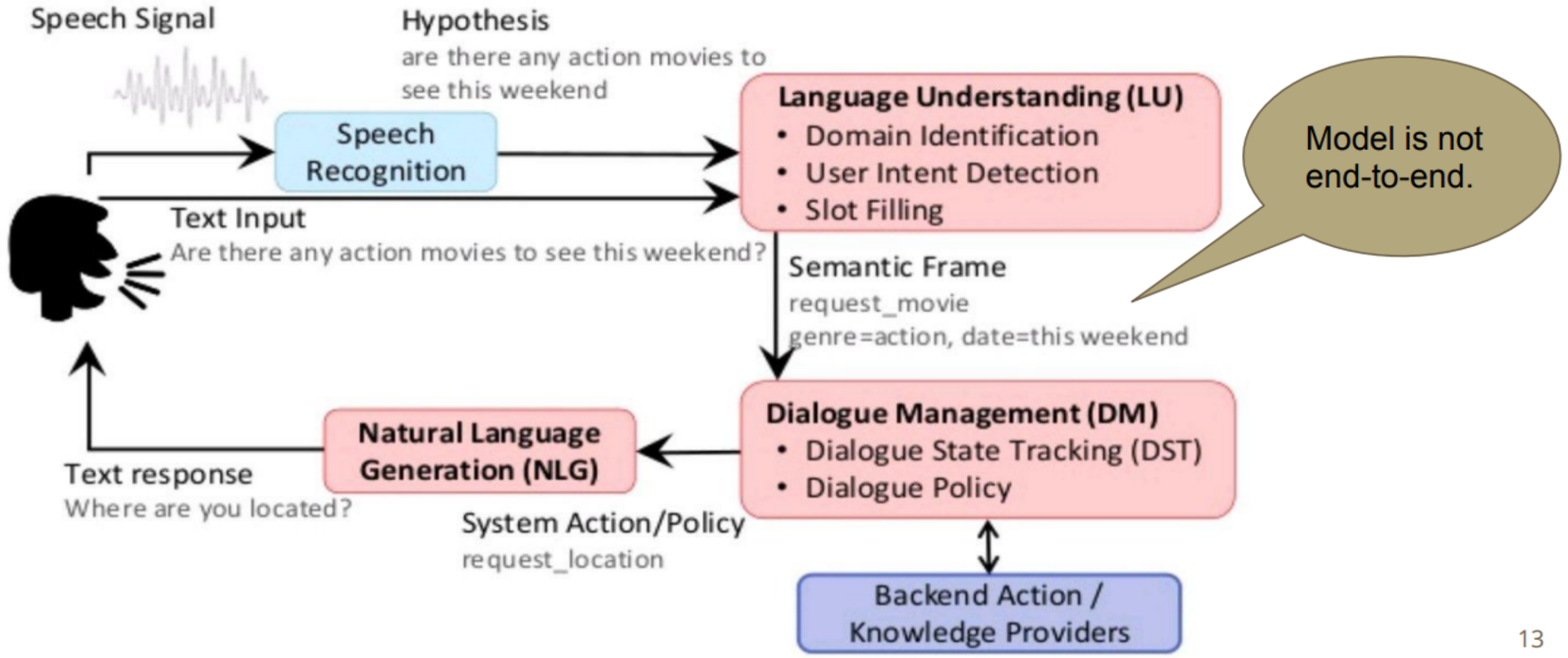
ZERO



GENERAL TOD SYSTEM

The dialogue state architecture

Task-Oriented Dialogue System (Young, 2000)

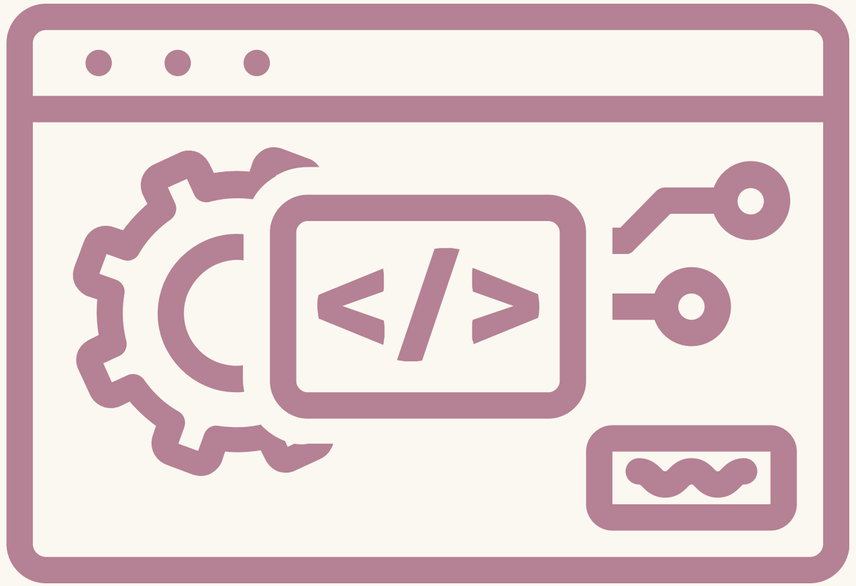


COMMON TASKS - ORDERED FROM START TO FINISH:

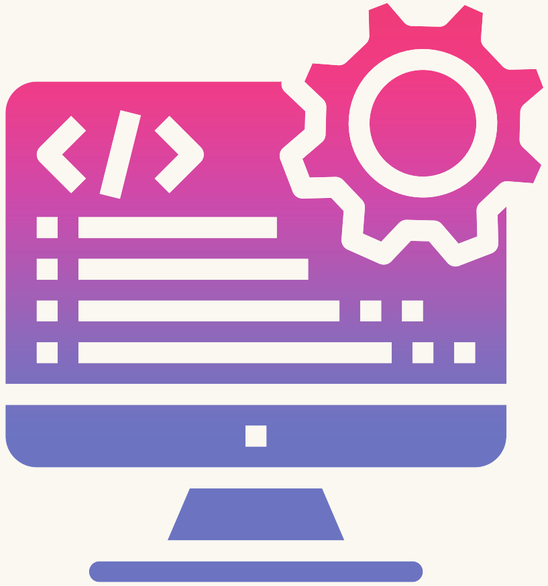
- **DOMAIN** CLASSIFICATION
- **INTENT** RECOGNITION
- **SLOT** FILLING
- **DIALOGUE STATE** TRACKING
- **DIALOGUE POLICY**
- RESPONSE GENERATION - OPTIONALLY SPEECH RECOGNITION

HTTPS://WWW.CS.PRINCETON.EDU/COURSES/ARCHIVE/SPRING20/COS598C/LECTURES/LEC16-TASK-ORIENTED-DIALOGUE.PDF

IN RECENT YEARS, MULTIPLE END-TO-END TASK-ORIENTED MODELS WERE PROPOSED



YET, THEIR CAPABILITIES ARE FULLY BASED ON PRE-TRAINED AND FINE-TUNING PROCESSES AND HARDLY ADAPT TO UNSEENED DOMAIN



FOR TOD SYSTEM, WE EMPHASIZE ON 2 ASPECTS:

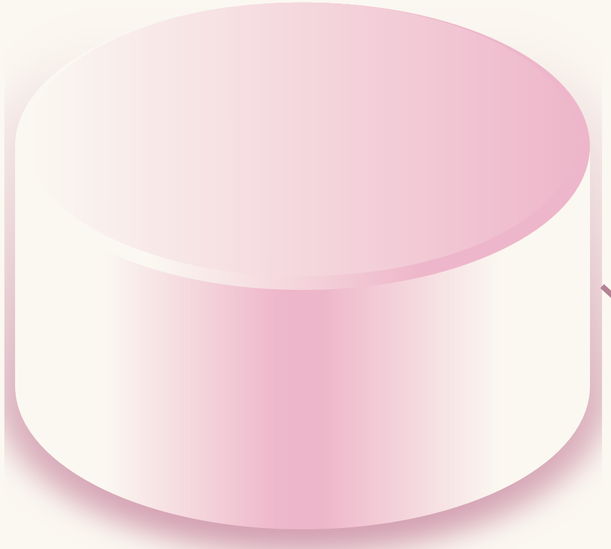
- GENERALITY: ABILITY TO SOLVE UNSEEN DOMAINS
- EXTENDABILITY: ABILITY TO ADAPT TO THE BIG SYSTEM

IN PREVIOUS RESEARCH, DIALOGUE FLOW WAS OMITTED

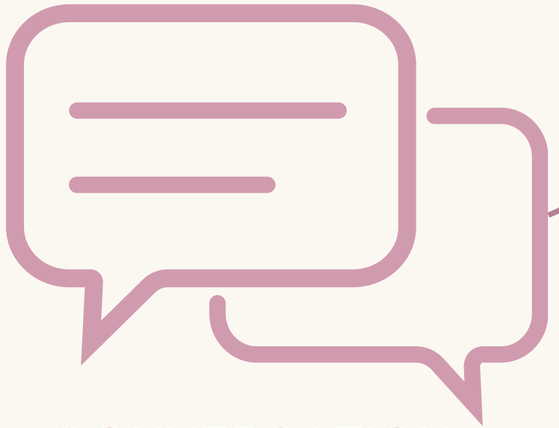
WE PROPOSE A METHOD FOR COMBINING SLOTS AND INTENTS RECOGNITION WHILE HIGHLY EXTENDABLE FOR FURTHER PROCESSING PHASES. EXTRA TAGS ARE ADDED AS CONTEXT FOR BETTER GENERALIZATION.

1. WHAT INFORMATION IS NEEDED TO MAKE A SHIP ORDER?
2. IN OUR ORGANIZATION, HOW SHOULD WE HANDLE SITUATIONS?

..... IN SHORT, WHAT IS THE DIALOGUE FLOW?



DATA SCHEMA

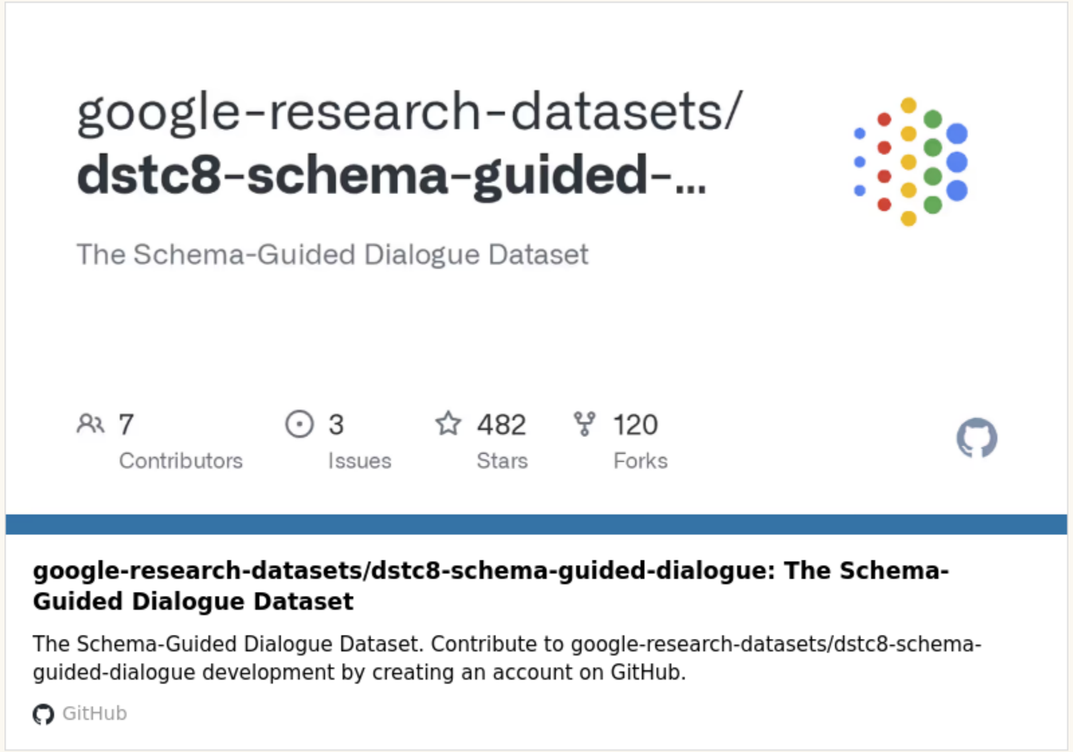


CONVERSATION

FOR A WORKER TO DO CUSTOMER SUPPORT

SGD

A schema-guided dialogue dataset is a collection of dialogues between a human and a virtual assistant, where the dialogues are annotated with a schema. The schema defines the set of intents and slots that can occur in the dialogue.



TERMINOLOGIES

USER TURN

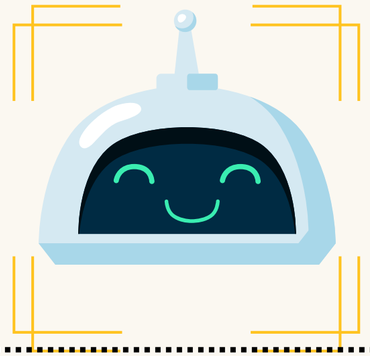
I AM TAKING A TRIP, CAN YOU HELP ME RESERVE MY TICKET PLEASE.

YES, I'M FLYING NEW YORK TO LAS VEGAS ON MONDAY NEXT WEEK. WHAT CAN YOU FIND?

WOULD I ARRIVE THE SAME DAY OR THE DAY AFTER?

IF I GO WITH THIS FLIGHT, WHERE WOULD I BE FLYING OUT OF AND WHERE WILL I ARRIVE?

SYSTEM TURN



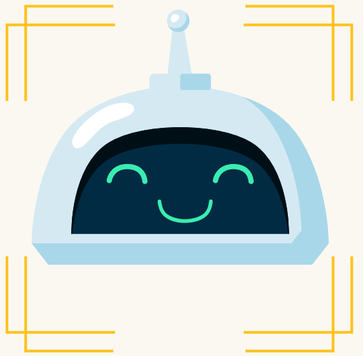
SURE, I WOULD LOVE TO. WHICH CITY WILL YOU BE VISITING? DO YOU KNOW WHAT CITY AND DATE YOU WILL BE DEPARTING?

IF YOU LIKE AMERICAN AIRLINES, THEY HAVE A 6:50 PM WITH 1 LAYOVER FOR $287 PER TICKET.

YOUR FLIGHT WILL DEPART AND ARRIVE ON THE SAME DATE.

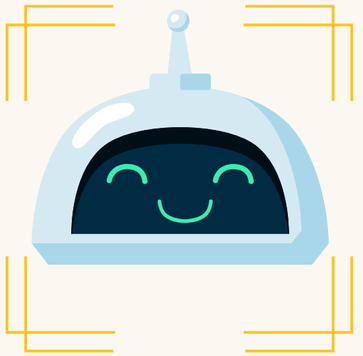
YOUR DEPARTURE AIRPORT IS JFK INTERNATIONAL AIRPORT AND ARRIVAL WOULD BE AT MCCARRAN AIRPORT.

TURN ANNOTATION EXAMPLE



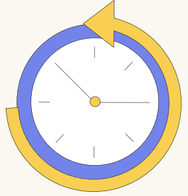

```
{
        "FRAMES": [
          {
            "ACTIONS": [
              {"ACT": "INFORM", "CANONICAL_VALUES": ["TRUE"],
                "SLOT": "PETS_ALLOWED",
                "VALUES": ["TRUE"]
              },
              {"ACT": "INFORM", "CANONICAL_VALUES": ["510-849-6628"],
                "SLOT": "PHONE_NUMBER",
                "VALUES": ["510-849-6628"]
              }
            ],
            "SERVICE": "HOMES_1",
            "SLOTS": [
              {"EXCLUSIVE_END": 34,
                "SLOT": "PHONE_NUMBER",
                "START": 22}
            ]
          }
        ],
        "SPEAKER": "SYSTEM",
        "UTTERANCE": "THE CONTACT NUMBER IS 510-849-6628 AND PETS ARE ALLOWED IN THE APARTMENT."
},
```

SCHEMA ANNOTATION EXAMPLE

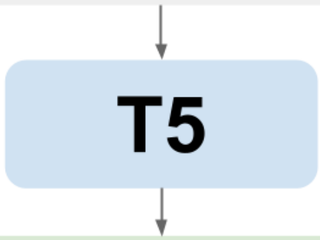


```
"SERVICE_NAME": "HOTELS_3",
"DESCRIPTION": "A LEADING PROVIDER FOR SEARCHING AND BOOKING HOTEL ROOMS",
"SLOTS": [
        { "NAME": "LOCATION", "DESCRIPTION": "LOCATION OF THE HOTEL"},
        { "NAME": "NUMBER_OF_ROOMS", "DESCRIPTION": "NUMBER OF ROOMS TO BOOK"},
        { "NAME": "CHECK_IN_DATE", "DESCRIPTION": "START DATE FOR THE HOTEL RESERVATION"},
        { "NAME": "CHECK_OUT_DATE", "DESCRIPTION": "END DATE FOR THE HOTEL RESERVATION"},
        ..................
],
"INTENTS": [
        { "NAME": "RESERVEHOTEL", "DESCRIPTION": "RESERVE A SELECTED HOTEL FOR GIVEN DATES",
          "REQUIRED_SLOTS": [ "HOTEL_NAME", "LOCATION", "CHECK_IN_DATE", "CHECK_OUT_DATE" ],
          "OPTIONAL_SLOTS": { "NUMBER_OF_ROOMS": "1" },
          "RESULT_SLOTS": [ "LOCATION", "NUMBER_OF_ROOMS", "CHECK_IN_DATE",
     "CHECK_OUT_DATE", "AVERAGE_RATING", "HOTEL_NAME", "STREET_ADDRESS",
     "PHONE_NUMBER", "PRICE", "PETS_WELCOME" ] },
        ..................
] },
POSSIBLE USER ACTIONS: [.....],
POSSIBLE SYSTEM ACTIONS: [.....],
```

PREVIOUS RESEARCH



[user] hey I am looking for a train from oxford to cambridge [agent] what date and time would you want to leave? [user] by 1pm on tuesday and getting there by 2pm [states] train-arriveby=2pm train-day=b of a) monday b) tuesday c) wednesday train-departby=1pm train-destination=cambridge train-departure=oxford [user] could you help me find a train to cambridge on wednesday? [agent] sure! what station would you like to leave from? and when would you like to depart? [user] london king's cross. i was wondering if there are any trains that arrive by 3pm.



[states] train-arriveby=3pm train-day=c train-departby=none train-destination=cambridge train-departure=london king's cross

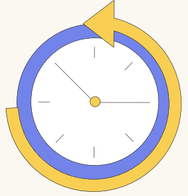
GOOGLE RESEARCHERS PUBLISHED 2 PAPERS:

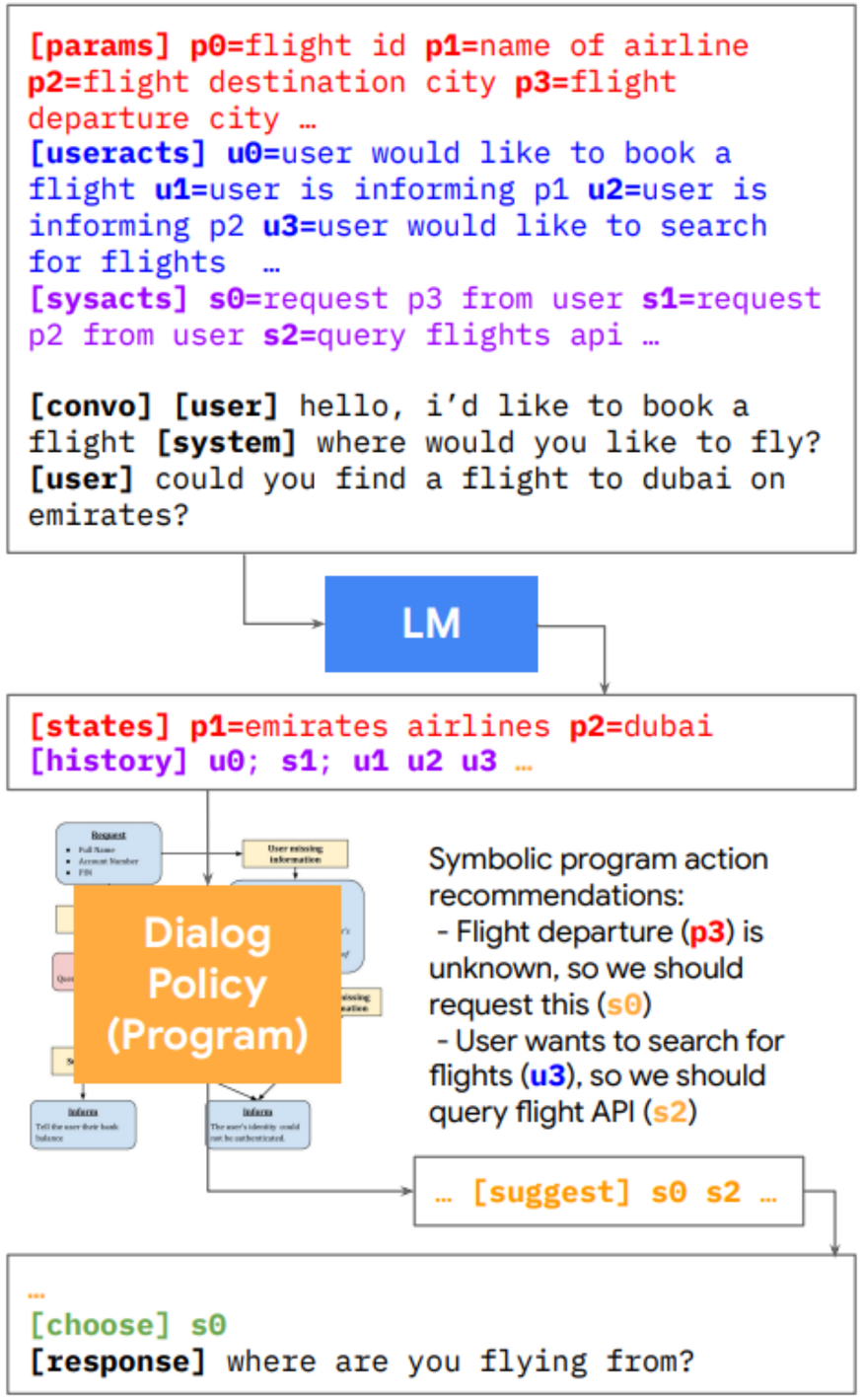
- DESCRIPTION-DRIVEN TASK-ORIENTED DIALOGUE
- SHOW DON'T TELL

INSTEAD OF TRADITIONAL CLASSIFICATION MODELS. THESE PAPERS MAKE USE OF LLMS TO TREAT STATE TRACKING AS MULTIPLE QAS.

HOWEVER, THE WEAKNESS IS THAT THEY DO NOT PROPOSE A FULL PROCESSING FLOW FOR TOD. (NOT EXTENDABLE)

PREVIOUS RESEARCH (CONT...)

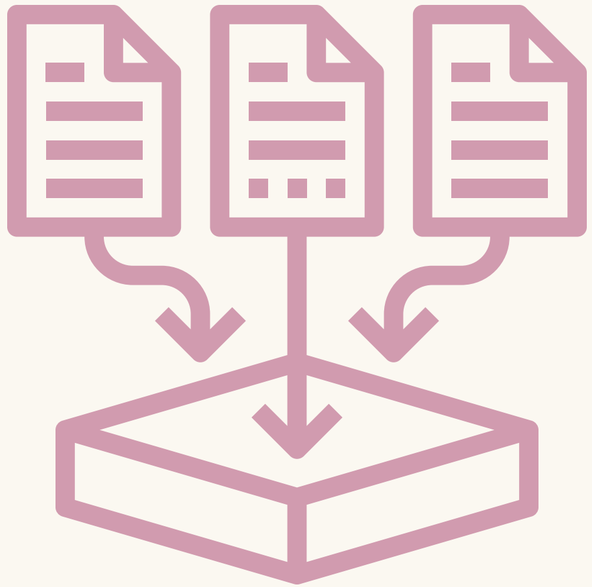


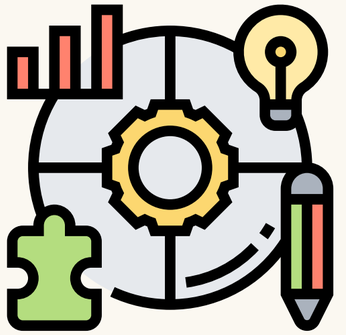


IN ANYTOD, BY EMBEDDING MORE SCHEMA CONTEXT INTO THE INPUT, AUTHORS ENHANCE EVEN FURTHER THE RESULT

THE AUTHORS SEPERATE ACTION (BEHAVIOR) AND SLOT (INFORMATION)

OUR METHOD





ADD MORE CONTEXT TO MODEL INPUT:

- GENERALIZE KNOWLEDGE ACROSS DOMAINS
- SOLVE DOMAIN TRANSITION, OPEN-DOMAIN TRANSITION

DIVERSE TASKS DURING TRAINING:

- MODEL CAPTURE BETTER AT STATE TRACKING
- SINGLE MODEL LEAD TO LESS RESOURCE WASTED

[**PARAMS**] P3=NUMBER OF SEATS TO RESERVE IN THE CAB; P63=EXPECTED WAITING TIME FOR PICK-UP BY CAB; P66=DESTINATION ADDRESS OR LOCATION FOR CAB; P93=TYPE OF CAB RIDE 93A) POOL 93B) REGULAR 93C) LUXURY; P173=TOTAL FARE FOR CAB RIDE

[**USERACTS**] U1=USER REQUEST UNDEFINED INFORMATION; U41=USER GOODBYE; U69=USER AGREE TO THE OFFER; U144=USER REQUEST P63; U148=USER REQUEST ALTERNATIVE ITEMS; U182=USER DENY THE OFFER; U203=USER REQUEST P173; U207=USER WANT TO GETRIDE; U274=USER INFORM UNDEFINED INFORMATION; U293=USER INFORM P93; U357=USER INFORM P66; U373=USER INFORM P3; U414=USER SELECT ITEM; U446=USER THANK

[**SYSACTS**] S11=INFORM P173; S87=ASK TO CONFIRM VALUE OF P93; S135=INFORM UNDEFINED INFORMATION; S208=ASK TO CONFIRM VALUE OF P3; S221=ASK TO CONFIRM UNDEFINED INFORMATION; S229=REQUEST UNDEFINED INFORMATION; S323=REQUEST P3; S353=GOODBYE USER; S394=QUERY GETRIDE API; S424=NOTIFY FAILURE; S468=INFORM NUMBER OF ITEMS SATIFIED USER; S472=ASK TO CONFIRM VALUE OF P66; S477=INFORM P63; S479=REQUEST P93; S485=REQUEST P66; S496=NOTIFY SUCCESS; S548=ASK USER IF THEY NEED ANYTHING MORE

[**DEPENDENCIES**] S87, S208, S472, U293, U357, U373 -> S394

[**TARGETACTS**]

[**CONVERSATION**] [USER] CAN YOU TELL ME WHAT TIME I CAN WATCH IT? [SYSTEM] SURE, WHAT DATE WILL YOU WATCH IT ON? [USER] IT'LL BE ON THE 8TH. [SYSTEM] OKAY, AND WHAT DO YOU WANT TO WATCH? AND WHERE? [USER] I'D LIKE TO WATCH CAPTAIN MARVEL IN PETALUMA.

...

[SYSTEM] IT'LL BE $31.40. [USER] THANKS SO MUCH. [SYSTEM] IS THERE ANYTHING ELSE? [USER] NO, THAT'S ALL THANKS.

21 TURNS UTTERANCE ARE LENGTHY, VARY AND LACK OF CONTEXT

NEW TAGS

POLICIES ARE EMBEDDED INTO INPUT, MORE EASIER FOR HUMAN AND POSSIBLY MACHINE

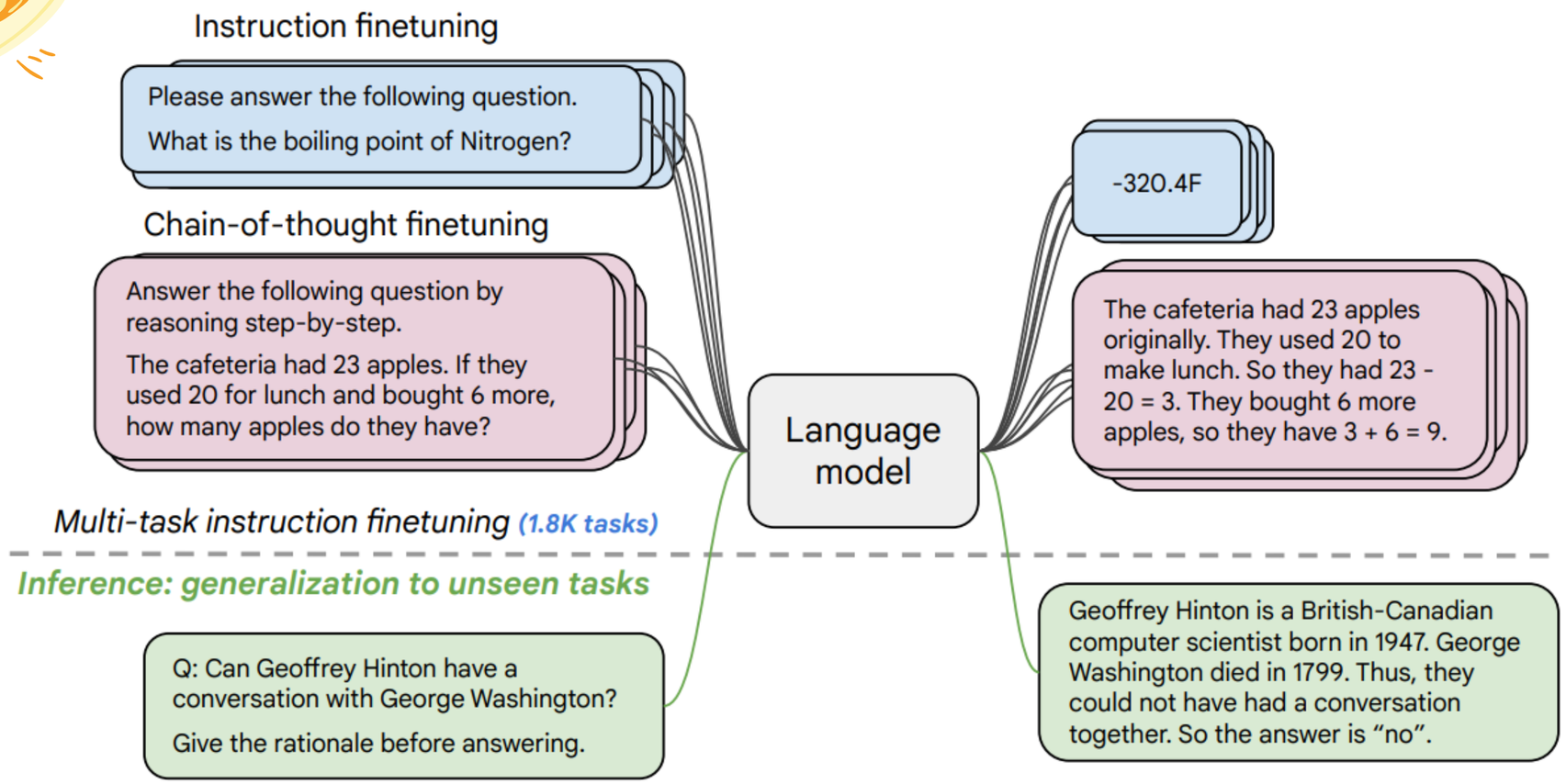
[**STATES**] P3=0507MQ2A; P66=0VB31; P93=LUXURY

[**HISTORY**] U274; S229; U274; S229, S229; U274, U274; S135, S135, S135; U1; S135; U1, U1; S135, S135; U207, U414; S323, S479; U293, U373; S87, S208, S472; U69; S394, S496; U203; S11; U446; S548; U182, U446

[**NEXTACTS**] S353

NEW TAGS

ADDING TOO MUCH CONTEXT, HOWEVER, MAKE INPUT LENGTHY AND HARD TO PROCESS, OUT SECOND EXPERIMENT FOCUS ON TASK BREAK DOWN AND MULTI-TASKS LEARNING



FLAN-T5 WAS TRAINED WITH INSTRUCTION FINETUNING

	INPUT	OUTPUT	METRICS
1	CONVERSATION PARAMS	STATE	JOINT GOAL ACCURACY FROM DSTC CONTEST
2	CONVERSATION PARAMS	HISTORY TURN	USER UNDEFINED ACTION F1 ACTION F1 FROM DSTC CONTEST
3	(USER | SYSTEM) ACTIONS		
4	PARAMS (USER + SYSTEM) ACTIONS DEPENDENCIES TARGET ACTIONS HISTORY	NEXT ACTIONS	SYSTEM ACTION F1 MANUAL DESIGN

Redesign TOD system

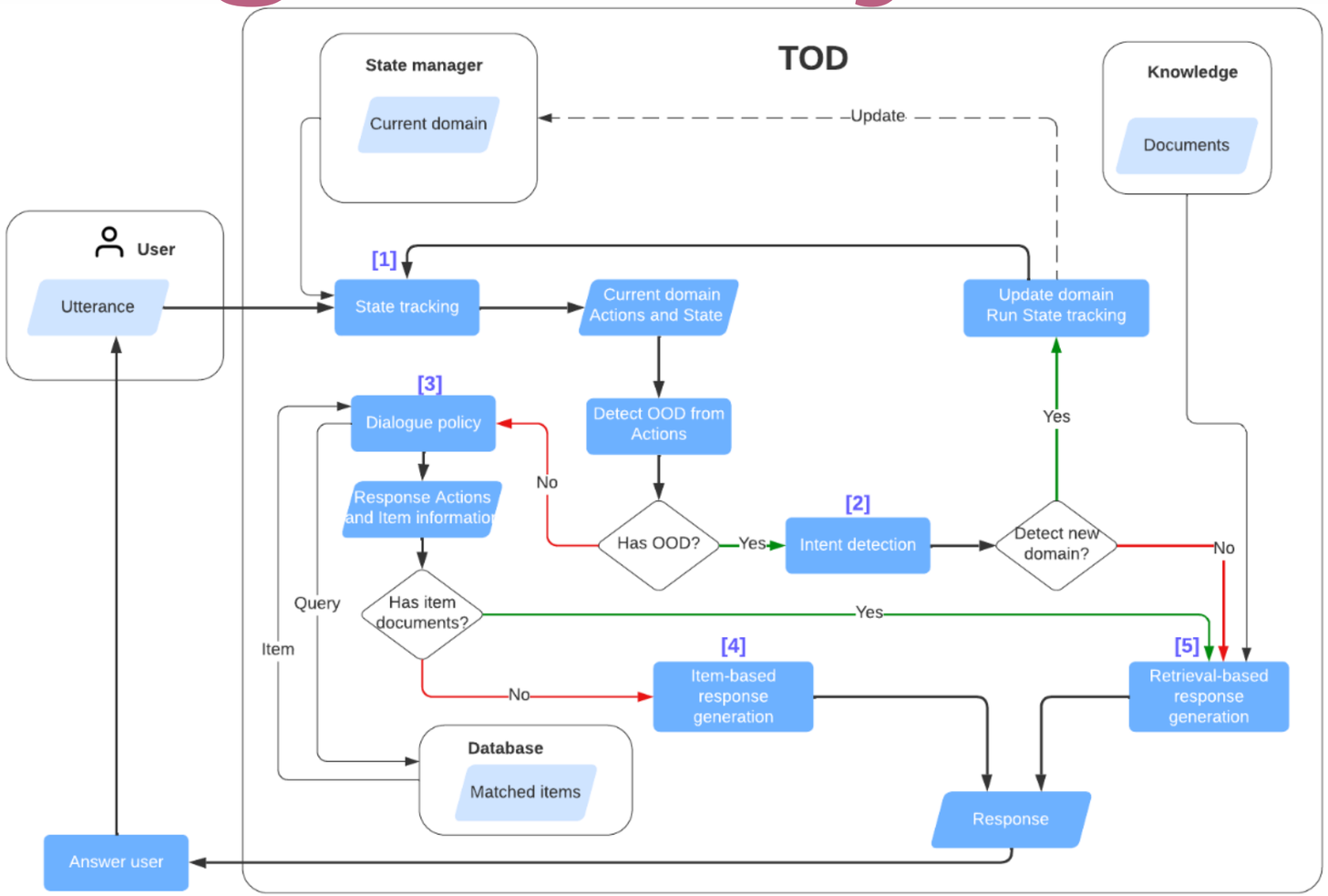


Figure 1: Overview of TOD system with retrieval capability. (1) capture user utterance to return abstracted actions, slot values and undefined actions. (2), triggered under undefined actions condition, to examine whether new intent is formed. In case exist new intent, we restart the flow with the new intent's schema. If the schema is not enough to response, (5) utilize document-based information to provide more accurate and knowledgeable response. Supposing state tracking (1) catch no abnormality, dialogue policy (3) use actions and slot values to interact with external storage. (3) supplies item-based information for (4) or document-based for (5).

Our contribution for DST

- ADDED UNDEFINED ACTIONS
  - ➡️ EASIER DOWNSTREAM
  - ➡️ PROXY FOR DOMAIN TRANSITION
- SYMBOLIZE SLOT VALUE
  - ➡️ BOOST JOINT GOAL ACCURACY
- TASK BREAKDOWN
  - ➡️ REDUCE TOTAL TRAINING TIME
  - ➡️ MERGE SLOT FILLING AND INTENT DETECTION
- EXAMPLE TOD FLOW

Results

Model	JGA Seen	JGA Unseen
SDT-seq T5 1.1 XXL	95.8	86.4
AT T5 1.1 Base	89.9	62.4
AT T5 1.1 XXL	94.8	82.2
$IFST_{Xtags\_Xrand}$ Base	77.9	61.1
- $Xtags$	78.4	63.9
- $Xtags\_Xrand$	77.2	60.7
$IFST_{Xtags\_Xrand}$ Large	85.4	72.2
- $Xtags$	86.6	75.0
- $Xtags\_Xrand$	85.1	72.3

TABLE I

JGA ON SGD TEST SET. RESULT OF AT AND SDT-SEQ ARE INFERRED FROM [23] [15]

https://www.overleaf.com/project/6535198968378c7f4b5ff462

Model	UUAF1	All AF1	All SaF1	Seen SaF1	Unseen
AT T5 1.1 Base	-	-	89.8	86.1	
AT T5 1.1 XXL	-	-	91.3	88.9	
$IFST_{Xtags\_Xrand}$ Base	85.7	66.3	85.9	82.1	
- $Xtags$	85.4	65.2	82.3	79.6	
- $Xtags\_Xrand$	85.6	65.2	81.9	79.5	
$IFST_{Xtags\_Xrand}$ Large	93.2	81.2	89.4	88.2	
- $Xtags$	85.7	66.3	85.9	82.1	
- $Xtags\_Xrand$	85.5	65.7	86.2	82.2	

TABLE II

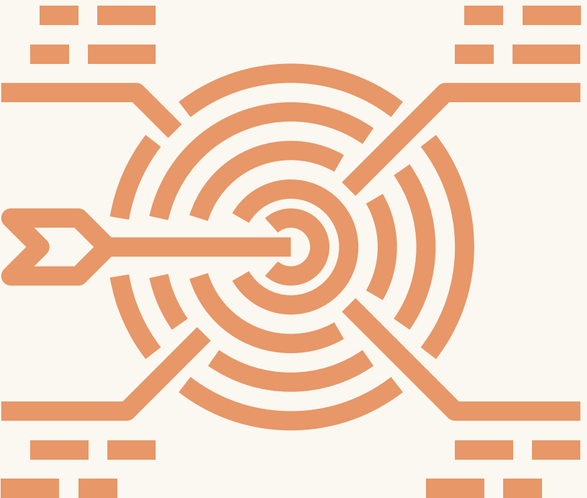
UUAF1, SAF1, AF1 ON SGD TEST SET. RESULTS OF AT ARE INFERRED FROM [23]. NOTICE SAF1 MEASURED IN IFST KEEPS TRACK OF SYSTEM QUERY AND INFORM ACTIONS ONLY

METRICS FOR DST ARE SIMPLE

AVG. SLOT ACCURACY: FRACTION OF SLOTS (AGGREGATED OVER ALL TURNS AND DIALOGUES) FOR WHICH THE MODEL PREDICTS THE CORRECT SLOT VALUE. - NP.MEAN

JOINT GOAL ACCURACY: FRACTION OF DIALOG TURNS FOR WHICH THE VALUES FOR ALL SLOTS THAT AREPREDICTED CORRECTLY. - NP.PROD

JGA IS MORE STRICT, ONLY 1 WRONG VALUE WOULD RESULT IN 0



EXAMPLE

PREDICTED: “HELLO WORKD”, “COIN”
LABEL: “HELLO WORLD”, “MONOPLY”

FUZZY SCORE: 0.9091, 0.3333
TURN AGA: 0.6212
TURN JGA: 0.3004

https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1204/reports/custom/report34.pdf

FORMULA

METRICS

levenshtein distance

FUZZY SCORE

$$\frac{(|a| + |b|) - lev_{a,b}(i,j)}{|a| + |b|}$$

JOINT GOAL ACCURACY: FRACTION OF DIALOG TURNS FOR WHICH THE VALUES FOR ALL SLOTS THAT ARE PREDICTED CORRECTLY.

$$outcome = \begin{cases} Positive & \text{for undefined\_action in turn} \\ Negative & \text{for other case} \end{cases} \quad (1)$$

$$outcome = \begin{cases} Positive & \text{for predict\_actions match label\_actions} \\ Negative & \text{for other case} \end{cases} \quad (2)$$

JOINT GOAL ACCURACY FROM DSTC CONTEST

USER UNDEFINED ACTION F1

ACTION F1 FROM DSTC CONTEST

SYSTEM ACTION F1 MANUAL DESIGN

FUTURE DIRECTIONS

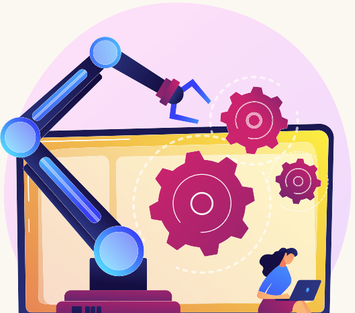
STATE TRACKING

ADD INTENT DETECTIONS ABILITIES

INCREASE OPEN DOMAIN CONTEXT

SYSTEM

COMPLETE THE SYSTEM



Appendix

TRANSITION: FLIGHT21

INPUT

[PARAMS] P7=THE TIME OF DEPARTURE FOR THE OUTBOUND FLIGHT; P16=CITY GOING TO; P20=LANDING TIME OF RETURN FLIGHT TO ORIGIN; P24=RETURN LEG FLIGHT DATE; P32=IF THE FLIGHT ARRIVES THE NEXT DAY AS PER DESTINATION LOCAL TIME ;...

[USERACTS] U2=USER INFORM UNDEFINED INFORMATION; ... U264=USER SELECT ITEM; U266=USER INFORM P133; U268=USER THANK

[SYSACTS] S1=INFORM P107; S23=INFORM P53; S24=INFORM P61; S27=INFORM P92; S54=QUERY SEARCHFORROUNDTRIPFLIGHT API; S61=OFFER USER P133; ...

[DEPENDENCIES] ... [TARGETACTS] ...

[CONVERSATION] [USER] I AM TAKING A TRIP, CAN YOU HELP ME RESERVE MY TICKET PLEASE. [SYSTEM] SURE, I WOULD LOVE TO. WHICH CITY WILL YOU BE VISITING? DO YOU KNOW WHAT CITY AND DATE YOU WILL BE DEPARTING FROM? ... [USER] I SEE. I WOULD LIKE TO HEAR MORE OPTIONS. WHAT DOES DELTA AIRLINES HAVE FOR ECONOMY SEATING PRICES? [SYSTEM] DELTA AIRLINES HAS 1 FLIGHT DEPARTING AT 3 PM WITH 1 LAYOVER FOR $328 PER PERSON. [USER] HOLD THAT THOUGHT, LET'S GO AND CHECK MY SAVINGS ACCOUNT BALANCE PLEASE.

OUTPUT

[STATES] P16=LAS VEGAS; P83=NEW YORK; P89=MONDAY NEXT WEEK; P107=ECONOMY; P133=DELTA AIRLINES [HISTORY] U248; S83, S106, S227; U82, U140, U145; S61, S190, S283, S312, S320; U220; S248; U75, U196; S88, S278; U111, U218, U266; S61, S190, S267, S283, S312, S320; U2, U2, U264 [NEXTACTS]

TRANSITION: BANK22

INPUT

[PARAMS] P9=HOW MANY DAYS THE TRANSFER WILL TAKE; P11=THE AMOUNT TO TRANSFER TO THE RECIPIENT; P34=TYPE OF USER'S BANK ACCOUNT: CHECKING OR SAVINGS 34A) CHECKING 34B) SAVINGS; ...

[USERACTS] U5=USER INFORM P11; U17=USER INFORM P34; U29=USER INFORM P114; U30=USER THANK; U40=USER DENY THE OFFER; U47=USER WANT TO GETACCOUNTBALANCE; U59=USER SELECT ITEM; ...

[SYSACTS] S14=REQUEST P11; S47=OFFER USER TRANSFERMONEYTOUSER; S50=OFFER USER P34; S53=REQUEST P34; ... S158=QUERY GETACCOUNTBALANCE API; ...

[DEPENDENCIES] S50, S100, S294, S314, U5, U17, U228 -> S88; U17 -> S158 [TARGETACTS] S158

[CONVERSATION] [USER] I AM TAKING A TRIP, CAN YOU HELP ME RESERVE MY TICKET PLEASE. [SYSTEM] SURE, I WOULD LOVE TO. WHICH CITY WILL YOU BE VISITING? DO YOU KNOW WHAT CITY AND DATE YOU WILL BE DEPARTING FROM? ... [USER] I SEE. I WOULD LIKE TO HEAR MORE OPTIONS. WHAT DOES DELTA AIRLINES HAVE FOR ECONOMY SEATING PRICES? [SYSTEM] DELTA AIRLINES HAS 1 FLIGHT DEPARTING AT 3 PM WITH 1 LAYOVER FOR $328 PER PERSON. [USER] HOLD THAT THOUGHT, LET'S GO AND CHECK MY SAVINGS ACCOUNT BALANCE PLEASE.

OUTPUT

[STATES] P34=SAVINGS [HISTORY] U207; S281, S281, S281; U207, U207, U207; S64, S64, S64, S64; U244; S64; U244, U244; S64, S64; U207, U207, U229; S64, S64, S64, S64, S64; U17, U47, U59; [NEXTACTS] S158 ;

MANAGEMENT TOOL