mForms: Multimodal Form-Filling with Question Answering

Larry Heck, Simon Heck, Anirudh Sundar Al Virtual Assistant (AVA) Lab Georgia Institute of Technology

Introduction *Slot-Filling* → *Form-Filling*







Early Work

Output layer for

0

each schema

/domain

Towards Zero-Shot Frame Semantic Parsing for Domain Scaling

ERSPEECH 2017 Situated interaction August 20-2

Form Filling with Concept Tagger

Ankur Bapna, Gokhan Tür, Dilek Hakkani-Tür, Larry Heck

Google Research, Mountain View





Multimodal Form-Filling



Deka, B.; Huang, Z.; Franzen, C.; Hibschman, J.; Afergan, D.; Li, Y.; Nichols, J.; and Kumar, R. *RICO: A Mobile App Dataset for Building Data-Driven Design Applications.* Proceedings of the 30th Annual Symposium on User Interface Software and Technology, UIST '17.

Traditional Question-Answering

In the 1960s, a series of discoveries, the most important of which was seafloor spreading, showed that the Earth's lithosphere, which includes the crust and rigid uppermost portion of the upper mantle, is separated into a number of tectonic plates that move across the plastically deforming, solid, upper mantle, which is called the asthenosphere. There is an intimate coupling between the movement of the plates on the surface and the convection of the mantle: oceanic plate motions and mantle convection currents always move in the same direction, because the oceanic lithosphere is the rigid upper thermal boundary layer of the convecting mantle. This coupling between rigid plates moving on the surface of the Earth and the convecting mantle is called plate tectonics.

What was the most important discovery that led to the understanding that Earth's lithosphere is separated into tectonic plates? seafloor spreading



Multimodal Form-Filling: Automatic GUI-to-Question Generation



Multimodal Form-Filling: Reformulated as Question Answering



Zero-Shot

Fine-Tuning

Approach Multimodal Form-Filling

Start/End Span

T,

C | T₁

T_N

...

T_[SEP]

T,'



Setup: Mobile Apps

f 🚺 🛈 🎇	8 🖻		🛡 🖹 🛿 1:43
Contraction Contra	e Logger		▶ ≣
	EN	TRY	HISTOR
TRIP	DAY	FUEL	OTHER
120000		Odometer Va	alue 🕜
		Description	n 💆 📎
		Veh	icle 🖌 🕂
✓ Track d ● E	istance with Business 〇	GPS Personal 〇 (Other
	Star	t trip	

F Search for Flight By City My Flights By Number Recent 9 Y Select departure airport 1Ť 9 Select arrival airport 1 U 11111 Monday, February 27, 2017 -Start new search







Corinthia Hotel London 2,271 Reviews #16 of 1,078 hotels in London

nightly from \$733*



Setup: Simulated Multimodal Form-Filling with ATIS Dataset

B-aircraft_code	what	is	the	aircraft code?
B-airline_code	what	is	the	airline code?
B-airline_name	what	is	the	airline name?
B-airport_code	what	is	the	airport code?
B-airport_name	what	is	the	airport name?
B-arrive_date.date_relative	what	is	the	relative arrival date?
B-arrive_date.day_name	what	is	the	day name of the arrival date?
B-arrive_date.day_number	what	is	the	day number of the arrival date?
B-arrive_date.month_name	what	is	the	month name of the arrival date?
B-arrive_date.today_relative	what	is	the	arrival date relative to today?
B-arrive_time.end_time	what	is	the	end time of the arrival?
B-arrive_time.period_mod	what	is	the	time period modification of the arrival?
B-arrive_time.period_of_day	what	is	the	period of day of the arrival time?
B-arrive_time.start_time	what	is	the	start time of the arrival?
B-arrive_time.time	what	is	the	arrival time?
B-arrive_time.time_relative	what	is	the	relative arrival time?
B-booking_class	what	is	the	booking class?
B-city_name	what	is	the	city name?
B-class_type	what	is	the	class type?
B-compartment	what	is	the	compartment

ATIS translated to a simulated form-filling task Slot types translated to Questions

Experiments Setup

Visual App	Slot descriptions
Vehicle Logger	fuel cost, fuel added, trip description, gps tracking, start logging, date, odometer value, trip type, entry, vehicle
United	arrival airport, departure airport, travel dates, search, switch/swap airports
ATIS form-filling	aircraft code, airline code, airline name, airport code, airport name, arrival date (relative), arrival date (day name), arrival date (day number), arrival date (month name), etc
Trip Advisor	number of beds, date range, filter by price, filter by rating, number of nights, number of people

Visual App	Sample Utterance	# Utterances
Vehicle Logger	"Please activate GPS tracking and log my car trip"	850
United	"Book a Flight from California to arizona on august 15th 2020"	850
ATIS form-filling	"I live in Denver and I'd like to make a trip to Pittsburgh"	4478
Trip Advisor	"Please book a 5 star hotel in Atlanta Georgia"	803

Accuracy of mForms vs "JB" (LMM-Based supervised slot filler**)



Effect of Visual Semantics in mForms: Vehicle Logger

# train samples	0	50	100	500
No Visuals	0.01	0.29	0.32	0.71
Text Visuals	0.36	0.69	0.71	0.88
GUI (all) Visuals	0.48	0.73	0.80	0.87

Effect of Multi-Task Training in mForms

# train samples	0	5	100	500
Vehicle Logger	0.48	0.46	0.80	0.87
+ATIS form-filling	0.52	0.60	0.80	0.89

Summary



- Robust to sparse training data conditions (scales well)
- mForms method yields SoTA accuracy on ATIS w/ 1/10th data



Start/End Span

T_M'

С Т,

T_N

T_[SEP]

mFORMS

T,'

mForms: Multimodal Form-Filling with Question Answering

Larry Heck, Simon Heck, Anirudh Sundar Al Virtual Assistant (AVA) Lab Georgia Institute of Technology

Thank you!