

Entrainment increases naturalness in dialogue systems, without hurting task success

LEEETs-Dial: Linguistic Entrainment in End-to-End Task-oriented Dialogue systems

Nalin Kumar
Ondřej Dušek

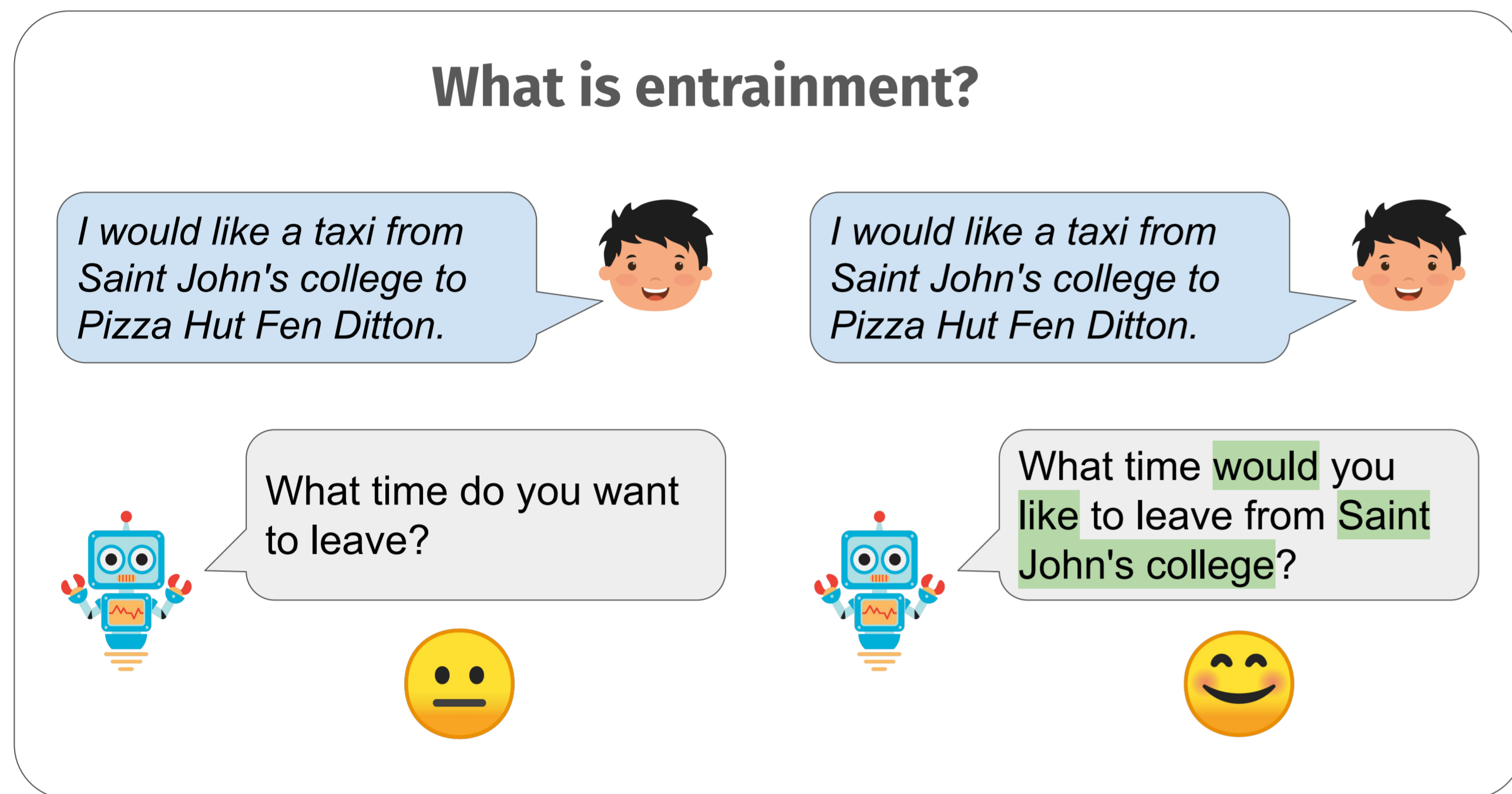
nalin.kumar775@student.cuni.cz
odusek@ufal.mff.cuni.cz



CHARLES
UNIVERSITY



What is entrainment?



How to facilitate entrainment? Ideas:

- + **IW**: Promote “nice” data instances during training
- + **ULL**: Maximize the probability of “desired” tokens
- + **LK**: Assist the model by adding hints to the input

Experiments & Results

Model: **AuGPT** (Kulhánek et al. 2021)

- end-to-end, GPT-2 based
- *context* → belief state → *DB* → response

Baselines

- **Base-CE**: AuGPT finetuned on MultiWOZ 2.1
- **D&J16**: Base-CE + n-gram reranking while decoding

Evaluation on MultiWOZ 2.1

- **MultiWOZ metrics**: inform/success rate & BLEU
- **Entrainment metrics**:
 - 1-gram precision & recall (lexical)
 - 2/3-gram precision on POS tags (syntactic)
- **Human**: Relative ranking of naturalness

Objective & Approaches

- **Lexical entrainment for E2E Task-oriented DS**
- Entrainment = alignment/adaptation of conversational partners
 - occurs at various linguistic levels
- Entrainment → **More natural** responses → **Better user experience**

I would like a taxi from Saint John's college to Pizza Hut Fen Ditton.

1. Instance Weighting (IW)

What time do you want to leave?

+1

What time would you like to leave from Saint John's college?

+10

- Higher weights for better user-system overlap (1-gram precision)
 - Discrete (**IW1**): 1 if 1-gram precision < 0.25, else 10
 - Continuous (**IW2**): modified sigmoid function with values 1-10

I would like a taxi...

2. User Likelihood Loss (ULL)

What time **do**...

Reuse {I, would, like, a, taxi, from, Saint, John's, college, to, Pizza, Hut, Fen, Ditton.}

- Maximize probability of user tokens $u \in U$
- Additional loss with weight α : $-\alpha \cdot \log(\sum_{u \in U} p(u|x_t))$

3. Conditioning on Lexical Keywords (LK)

I would like a taxi...

Use these keywords: {would, like, taxi, from, Saint, John's, college}

What time would you like the taxi to leave from Saint John's college?

- Assisting model with “important” user tokens
 - Training keywords = overlapping 1-grams in user-system
 - Inference keywords = user tokens based on self-attn scores
- Blending: expose model to inference keywords during training with probability $\sigma = \{0, 0.05, 0.5\}$

Results

system	inform	success	bleu	lex-p1	lex-r1	syn-p2	syn-p3	Ørank
Base-CE	83.5	63.8	15.7	20.7	24.5	14.8	5.0	4.18
D&J16	85.7	63.6	10.6	31.9	26.1	23.1	10.4	5.35
IW1	84.5	68.6	14.9	22.9	30.9	16.4	5.9	3.16
IW2	82.6	67.7	15.3	22.9	29.8	16.4	5.8	3.77
ULL (0.20)	81.6	65.3	15.3	23.7	29.4	16.2	5.7	4.17
ULL (0.25)	81.6	63.6	14.6	24.7	31.6	16.9	6.1	4.33
LK (0.05)	83.3	66.3	12.8	25.8	33.6	17.0	6.5	3.25

- **D&J16** – shorter, less polite, less interactive outputs
- **IW1** – consistently natural-looking outputs
- **LK** – better entrained outputs
- **ULL** – fragile, hallucinates with higher weight

Example

Context: [...]

User: I would like one in the centre please.

Ground Truth: Christ's College, Corpus Christi, King's College, and Pembroke College are some of the colleges in that area

Base-CE: Christ's College is located in the centre and has free admission

D&J16: Christ's College is in the centre of town.

IW1: There are 13 colleges in the centre. I recommend Christ's College. Would you like more information on it?

ULL (0.25): I would recommend Christ's College. Would you like the address?

LK-CE (0.05): Christ's College is one of the best college in the centre. Would you like more information?



Presented at NAACL 2024, Mexico City

<https://github.com/knalin55/LEEETs-Dial>

Funded by the European Union (ERC NG-NLG 101039303 & H2020 Humane-AI-Net 952026).

Using LINDAT/CLARIAH-CZ Research Infrastructure (Czech Ministry of Education, Youth and Sports project No. LM2018101).

