Target-Guided Open-Domain Conversation Planning

Yosuke Kishinami[1], Reina Akama[1,2], Shiki Sato[1], Ryoko Tokuhisa[1] Jun Suzuki[1,2], Kentaro Inui[1,2]
[1] Tohoku University [2] RIKEN
yosuke.kishinami.q8@dc.tohoku.ac.jp

Overview

- We proposed **task to evaluate conversation planning abilities** of conversational agents
- Experiments reveal **keyword sequence planning** improves agent's planning ability

Background: Planning in target-oriented conversation

Many studies address **target-oriented tasks** to realize variety of applications [Tang+'19;Qin+'20;Zhong+'21]
(e.g., therapeutic conversation, recommendation, …)

e.g., **Target-Guided Open-Domain Conversation (TGC)** [Tang+'19]

Agent's target: leading conversation to mention target word

Target word: ***pizza***

How are you?
I'm good. Thank you for asking.
*Have you had **lunch**?*
I like Chinese food.
*Nice! I like Italian. Especially **pizza**.*

This task setting is similar to **AI Planning**
(Both settings are given **initial state** and **target** to be achieved)

Could Planning be applied to TGC?

Our idea: Evaluate target-oriented agent's planning ability by self-conversation

First step: analysis planning ability of current conversational agents

Target word: ***book***

How are you?
I'm fine. Thanks.
*I often read **books** in that series.*

Generate self-conversation (Multiple-step plan)

Evaluate agent's planning ability

Proposed task: Target-Guided Open-Domain Conversation Planning (TGCP)

Task overview

Input
Initial utterance: *Hi, what do you do for living?*
Target word: ***book***

Conversational Agent

Output (conversation plan to achieve target)
Hi, what do you do for living?
I work as an engineer.
That sounds nice. How do you learn coding?
*I read and learn from technical **books**.*

Evaluation metrics

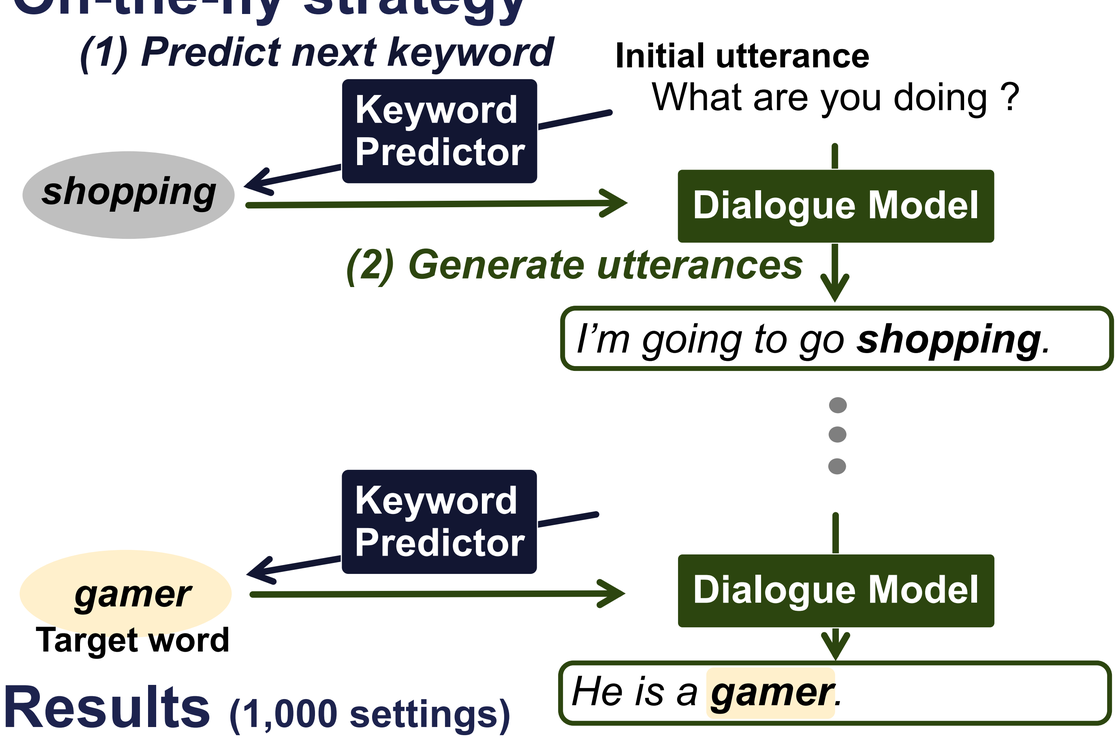
- **Achievement** *Whether target word is mentioned?*
- **Smoothness** *Whether utterance transitions is natural?* (5-point Likert-scale)

Generate conversation with only agent
Evaluate agent's planning ability without dialogue partners (e.g., Human)

Experiments: Investigate of planning abilities of dialogue models with strategy

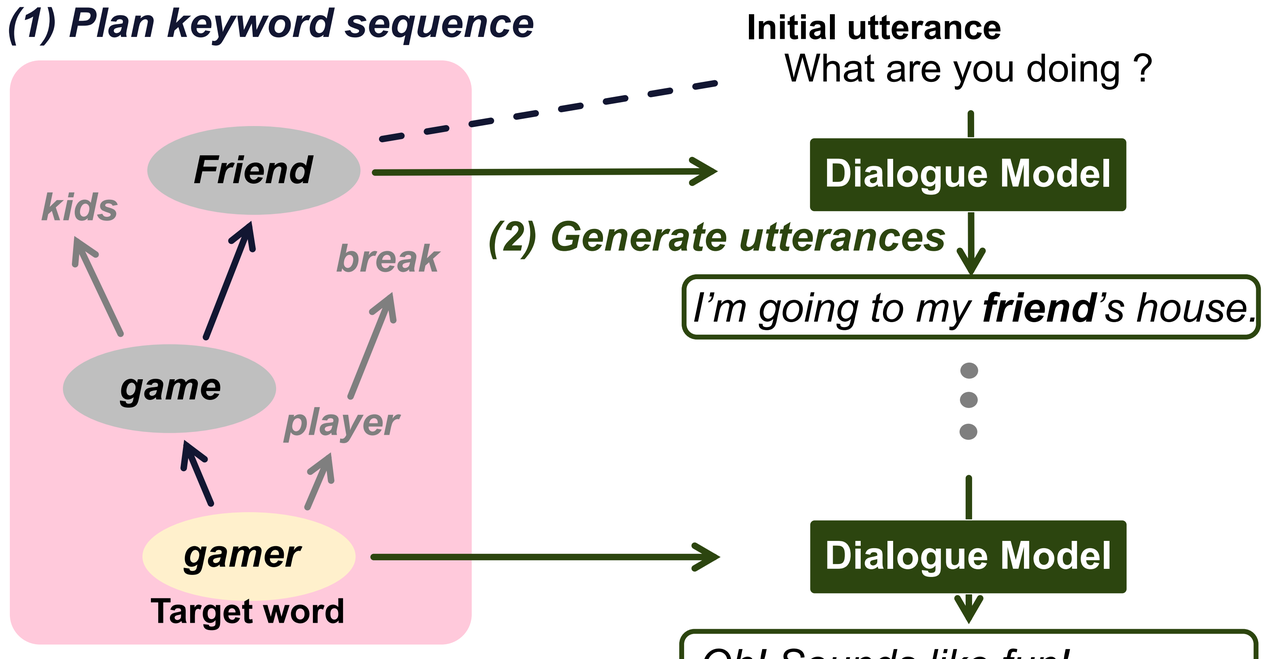
On-the-fly strategy

(1) Predict next keyword



Pre-design strategy

(1) Plan keyword sequence



Results (1,000 settings)

Models	Strategy	Achievement	Smoothness
Blender[Roller+'21]+CKC[Zhong+'21]	On-the-fly	0.247	3.90
Blender[Roller+'21]+PreDes.	Pre-design	**0.425**	4.05
Human	-	1.000	4.11

(5:Strongly Good ~ 1:Strongly Bad)

Pre-design strategy **improves achievement**