

発話文	$\sum_{o \in O} F_{S, r_{other}}(o)$	
ハンバーグの中で、まぐろを使っている料理もあるんだけど知ってる？	15	○
ワインの中で、温度があたたかいものがあるんだけど知ってる？	21	○
コーヒーで味がほろ苦い料理って知ってる？	5	×
うどんで主食として食べられる料理って知ってる？	29	×

表 1: ニュース性のある情報を含むものとして出力された発話の例

発話文	$\sum_{o \in O} F_{S, r_{other}}(o)$	
米を使っている料理で料理種が和食の料理多いよね	315	○
肉料理でおかずとして食べる料理多いよね	78	○
鶏肉を使っている料理で料理種が和食の料理多いよね	59	×
野菜類で沖縄でよく食べられているよね	3	×

表 2: 一般性のある情報を含むものとして出力された発話の例

のある情報として出力されたものである。知識グラフ中のエンティティの集合 $S = \{s | (s, IS_A, \text{ハンバーグ}) \text{ exists in GDB} \}$ の多くが、材料として「ひき肉」や「玉ねぎ」を使用している中で、「まぐろ」を使用するものもあるという情報を抽出した結果である。

全体的な出力の傾向として、 $\sum_{o \in O} F_{S, r_{other}}(o)$ の数が少ない時には特徴的な情報、特に一般性のある情報が得られない事例が多いことが分かった。原因として、GDB中に当該集合に属するエンティティが十分にはなく、現実の分布と対応していないことが原因だと想定できる。このような情報に基づき得られた発話は採用しないことで、得られる発話候補集合に不適切なものが混入するのを防ぐ必要がある。

4 おわりに

本稿では知識グラフから、エンティティの集合に関する特徴的な情報を抽出する手法と、現状の出力結果を示した。エントロピーを利用した抽出方法を説明したが、この点については現在検討を進めている。

現状の問題点として、知識グラフに含まれるデータの量の差を考慮していないことが挙げられる。知識グラフは人手で構築しているため、密な部分と疎な部分が存在する。例えば、料理 DB の中には $(s, \text{CUISINE_TYPE}, \text{和食})$ となるような s が多く含まれている。このため、和食の生起頻度の高さが分布に影響を与え、表 2 の上から 3 番目のような発話得られることがある。今後、この点を考慮した知識の抽出法の改良を検討する。

今後の他の方向性として、知識グラフ補完結果の利用がある。知識グラフ補完結果により知識グラフを部分的に拡張することで、より妥当な知識の抽出を図る。また本研究で構築した手法が、対象とする知識グラフに依存するか否かを確認するために、今回用いた知識グラフ以外も使用して結果を検証する。

参考文献

- [1] Oriol Vinyals and Quoc V. Le. A neural conversational model. *CoRR*, abs/1506.05869, 2015.
- [2] Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky. Adversarial learning for neural dialogue generation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 2157–2169, 2017.
- [3] Iulian V. Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. Building end-to-end dialogue systems using generative hierarchical neural network models. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, page 3776–3783, 2016.
- [4] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A diversity-promoting objective function for neural conversation models. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 110–119, 2016.
- [5] Mikio Nakano and Kazunori Komatani. A framework for building closed-domain chat dialogue systems. *Knowledge-Based Systems*, 204:106212, 2020.