1 We thank all the reviewers for their insightful comments, suggestions, and positive feedback. We will update the paper
2 based on the constructive comments provided. We start by addressing a common point regarding the use of OpenIE
3 methods for text-based games, and then address questions from individual reviewers.

4 **Comparison with OpenIE-based agents.** One common point raised in the reviews is the relationship between our
5 approach and approaches that use OpenIE algorithms for graph construction, such as KG-DQN [4] and KG-A2C [3].
6 These are indeed important prior works, and the key distinction between GATA (our work) and [4] and [3] is that they
7 rely on handcrafted filters combined with OpenIE algorithms, whereas GATA is entirely data-driven. We actually put
8 significant effort into including [4] as a baseline in the months prior to our submission. However, unfortunately, we
9 were unable to faithfully reproduce their model or achieve non-trivial performance on our evaluation games. These
10 problems persisted even after contacting the original authors for assistance, and we believe two key issues were (a)
11 the hand-crafted filters in [4] not generalizing across games and (b) substantial scalability issues. More generally,
12 off-the-shelf OpenIE tools are suboptimal for maintaining dynamic knowledge graphs in text-based games; in particular:
13 • Low precision: OpenIE generates a large amount of unrelated tuples (e.g., from a sentence "what is that?" it extracts
14 nodes "what" and "that" connected by edge "is"). Empirically, a set of filtering heuristics are needed as used in [3, 4].
15 • Low recall: Tools like OpenIE can hardly cover certain node types appearing in text-based games. Some nodes that
16 describe states of objects are absent from O_t. For instance, the resulting observation of making a meal will be "you
17 made a meal", but the "consumed" states of the ingredients are absent from O_t thus require reasoning. Typically, such
18 types of nodes are not entities (nor even noun phrases), making it difficult for tools like OpenIE to detect them.
19 • Dynamicity: Since we are dealing with a dynamic environment, some handcrafted heuristics for maintaining/updating
20 the knowledge graph (tuples extracted by OpenIE) are needed. For instance, after moving to a new room, the content of
21 the KG shouldn't be discarded but some relations need to be changed (e.g., location of the player).

22 **Reviewer #1:**
23 Analysing generated graphs: R1 raises an important point on analysing the graphs generated. For the analyses of
24 information encoded by the graphs, we refer them to Appendix D.5. Here, apart from the high-res visualization of the
25 graphs, we conduct probing tasks to study the information encoded by the generated belief graphs. Results suggest the
26 graphs can encode important information (Table 7). We will further discuss these results in the revision.
27 Other games: An alternative testbed would be to use the curated list of text-based games supported by Jericho [14].
28 However, the list is small with only 50 highly diverse games. Attempting to solve the very difficult problem of
29 out-of-distribution generalization (OODG) on diverse Jericho games was out-of-scope for this current submission,
30 which focuses on achieving state-of-the-art results on in-domain generalization. However, improving on GATA to
31 tackle Jericho and other more challenging generalization tasks is an objective of future work.

32 **Reviewer #2:**
33 Motivation for COC: Both the graph updater pre-training approaches (OG and COC) aim to extract useful information
34 that is sufficient to reconstruct the text observation. Contrastive learning has been shown to be an effective unsupervised
35 training regime to learn useful information in several prior works [7, 43, 22]. The motivation behind contrastive
36 unsupervised training is that one does not require to train complex decoders. For example, compared to OG, the COC's
37 objective relaxes the need for learning syntactical or grammatical features and allows GATA to focus on learning the
38 semantics of the O_t. This is an important point, and we will clarify this motivation in the revised main text.
39 Graph as memory: We agree with the reviewer that a graph-based representation also serves as a structured memory for
40 the agent. In particular, we believe such structured representations (KG) generalize well for text-based games owing to
41 two reasons: a) they serve as a good inductive bias as the states can be factorized in terms of nodes and relations; and b)
42 they act as a memory which aids the agent to act effectively in a partially observed scenario.

43 **Reviewer #3:**
44 Scalability: You raise some important clarification points regarding scalability. One important clarification is that at
45 every step, the belief graph $\mathcal{G} \in [-1, 1]^{\mathcal{R} \times \mathcal{N} \times \mathcal{N}}$ is decoded using f_d from a single vector h_t (as illustrated in Fig 2), and
46 the size of this vector h_t is significantly smaller than the full tensor (i.e., dimension 64, as described in appendices). We
47 use this low-dimensional vector as the model's recurrent state, which alleviates scalability issues during learning. More
48 generally, regarding scalability and relationships to traditional knowledge base construction (KBC), it is important to
49 note that our approach is more related to recent work on neural relational inference (NRI, Kipf et al., arXiv:1802.04687)
50 than traditional KBC; in particular, we seek to generate task-specific graphs, which tend to be dynamic, contextual and
51 relatively small, whereas traditional KBC focus on generating large, static graphs. As a result, we believe discussing
52 extrapolations of GATA's graph extraction to IE domains is out of scope of this paper.
53 OpenIE: Please see our general response. Thank you for pointing this out and providing the references, we acknowledge
54 that our approach is related to information extraction literature and we will discuss them in our revision.

55 **Reviewer #4:**
56 Table 2: We will make the presentation of our experiments more clear in the revised version.