AUTOMATED DATA CARTRIDGE FOR CONVERSATIONAL AI BOTS

TECHNICAL FIELD

[0001] This disclosure relates to artificially intelligent virtual agents and, in particular, to computer-generated testing of artificially intelligent virtual agents.

BACKGROUND

[0002] Virtual agents, such as chatbots, emailbots, salesbots, and/or AI assistants, may be deployed with intelligent algorithms working in the background to assist users with inquiries and provide suggestions, directions, and/or other useful information. When users are interacting with the virtual agent, the virtual agent interprets and responds to natural language utterances. A virtual agent under test may be trained with various conversational scenarios to increase the accuracy and authenticity of the virtual agent. Present approaches to conversationally testing and training virtual agents involve manual effort to generate test data with limited test coverage.

SUMMARY OF THE INVENTION

[0002a] According to a first aspect of the invention, there is provided a system comprising: a processor, the processor configured to: obtain a test utterance sample from test utterance samples configured to test a virtual agent; tokenize the test utterance sample into sample tokens; determine, based on a natural language processing framework, a synonymous set (synset) for at least one of the sample tokens, the synset comprising lemma tokens that are synonymous with the at least one of the sample tokens; obtain a contextual corpus comprising corpus tokens; select, from the lemma tokens, a subset of lemma tokens that are included in the corpus tokens; generate a plurality of new test utterances; each of the new test

utterances having a same intent or meaning or both same intent and same meaning as the test utterance sample and comprising one or more lemma tokens from the subset of lemma tokens; generate, based on the natural language processing framework, a semantic relation metric indicative of semantic correspondence between the plurality of new test utterances and the test utterance sample; determine whether the semantic relation metric is greater than a predetermined threshold value for each of the plurality of new test utterances; insert, in response to the semantic relation metric being greater than the predetermined threshold value, a set of the new test utterances in a test repository; and communicate the set of the new test utterances to the virtual agent to test a response by the virtual agent to each new test utterance in the set of the new test utterances.

[0002b] According to a second aspect of the invention, there is provided a method, comprising: obtaining a test utterance sample from test utterance samples configured to test a virtual agent; tokenizing the test utterance sample into sample tokens; determining, based on a natural language processing framework, a synonymous set (synset) for at least one of the sample tokens, the synset comprising lemma tokens that are synonymous with the at least one of the sample tokens; obtaining a contextual corpus comprising corpus tokens; selecting, from the lemma tokens, a subset of lemma tokens that are included in the corpus tokens; generating a plurality of new test utterances, each of the plurality of new test utterances comprising at least one lemma token included in the subset of lemma tokens; generating, based on the natural language processing framework, a semantic relation metric indicative of semantic correspondence between the new test utterances and the test utterance sample; determining the semantic relation metric is greater than a threshold value for only a portion of the new test utterances; inserting, in response to the semantic relation metric being greater than the threshold value, the portion of the new test utterances in a test repository; and testing the virtual agent based on the portion of the new test utterances.

[0002c] According to a third aspect of the invention, there is provided a non-transitory computer readable storage medium, comprising: instructions executable by a processor, the instructions configured to cause the processor to: obtain a test

utterance sample configured to test a virtual agent; tokenize the test utterance sample into sample tokens; determine, based on a natural language processing framework, a synonymous set (synset) for at least one of the sample tokens, the synset comprising lemma tokens that are synonymous with the at least one of the sample tokens; obtain a contextual corpus comprising corpus tokens; select, from the lemma tokens, a subset of the lemma tokens that are included in the corpus tokens; generate a plurality of new test utterances, each of the new test utterances comprising at least one of the lemma tokens included in the subset; generate, based on the natural language processing framework, a respective semantic relation metric for each of the new test utterances, the respective semantic relation metric indicative of semantic correspondence between each of the new test utterances and the test utterance sample; determine the semantic relation metric is greater than a predetermined similarity threshold value for similar new test utterances, the similar new test utterances being identified as similar to the test utterance sample by the predetermined similarity threshold value; insert, in response to the semantic relation metric being greater than the predetermined similarity threshold value, the similar new test utterances in a test repository; and communicate the similar new test utterances to the virtual agent to test the virtual agent based on the similar new test utterances.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The embodiments may be better understood with reference to the following drawings and description. The components in the figures are not necessarily to scale. Moreover, in the figures, like-referenced numerals designate corresponding parts throughout the different views.

[0004] FIG. 1 illustrates an example of a natural language test compilation system;

[0005] FIG. 2 illustrates an example of the virtual agent test generation pipeline for a system;

CLAIMS

What is claimed is:

1. A system comprising:

a processor, the processor configured to:

obtain a test utterance sample from test utterance samples configured to test a virtual agent;

tokenize the test utterance sample into sample tokens;

determine, based on a natural language processing framework, a synonymous set (synset) for at least one of the sample tokens, the synset comprising lemma tokens that are synonymous with the at least one of the sample tokens;

obtain a contextual corpus comprising corpus tokens;

select, from the lemma tokens, a subset of lemma tokens that are included in the corpus tokens;

generate a plurality of new test utterances; each of the new test utterances having a same intent or meaning or both same intent and same meaning as the test utterance sample and comprising one or more lemma tokens from the subset of lemma tokens;

generate, based on the natural language processing framework, a semantic relation metric indicative of semantic correspondence between the plurality of new test utterances and the test utterance sample;

determine whether the semantic relation metric is greater than a predetermined threshold value for each of the plurality of new test utterances;

insert, in response to the semantic relation metric being greater than the predetermined threshold value, a set of the new test utterances in a test repository; and

communicate the set of the new test utterances to the virtual agent to test a response by the virtual agent to each new test utterance in the set of the new test utterances.

2. The system of claim 1, wherein to determine, based on the natural language processing framework, the synset for at least one of the sample tokens, the processor is further configured to:

tag each of the sample tokens with a corresponding part of speech (POS) tag;

identify, from the tagged sample tokens, at least one of the sample tokens tagged with the corresponding POS tag that is indicative of a stop word;

filter the sample tokens by removing the at least one of the sample tokens tagged with the corresponding POS tag indicative of the stop word; and

determine, based on the natural language processing framework, the synset based on at least one remaining sample token of the filtered sample tokens.

3. The system of claim 1, wherein to select, from the lemma tokens, the subset of lemma tokens that are included in corpus tokens, the processor is further configured to:

generate a frequency distribution of the corpus tokens; and

select, from the lemma tokens, the subset of lemma tokens in response to the lemma tokens in the subset of lemma tokens corresponding to corpus tokens above a threshold in the frequency distribution.

4. The system of claim 1, wherein to generate the plurality of new test utterances, the processor is further configured to:

generate a frequency distribution of the corpus tokens;

identify at least one of the sample tokens that corresponds to at least one of the corpus tokens;

select, from the at least one of the sample tokens, a sample token that corresponds to a corpus token that is associated with a frequency value that is greater than a threshold value; and

replace the selected sample token with a lemma token from the subset of lemma tokens.

5. The system of claim 1, further comprising:

receive a plurality of respective test results in response to communicating the set of the new test utterances to the virtual agent;

determine that the virtual agent failed a test based on a comparison of the test results with expected test results; and

retrain a trained corpus for the virtual agent based on the set of the new test utterances.

6. The system of claim 1, wherein the processor is further configured to:

receive a context identifier associated with the contextual corpus;

select computer-generated test utterances based on the contextual corpus;

communicate the computer-generated test utterances to the virtual agent;

receive test results from the virtual agent;

compare the test results with expected test results; and

flag, in response to the received test results not corresponding to the expected test results, the computer-generated test utterances in the test repository.

7. The system of claim 1, wherein to obtain the contextual corpus, the processor is further configured to:

receive a web address corresponding to a source of the contextual corpus; and

access the contextual corpus from a remote server based on the web address.

8. A method, comprising:

obtaining a test utterance sample from test utterance samples configured to test a virtual agent;

tokenizing the test utterance sample into sample tokens;

determining, based on a natural language processing framework, a synonymous set (synset) for at least one of the sample tokens, the synset comprising lemma tokens that are synonymous with the at least one of the sample tokens;

obtaining a contextual corpus comprising corpus tokens;

selecting, from the lemma tokens, a subset of lemma tokens that are included in the corpus tokens;

generating a plurality of new test utterances, each of the plurality of new test utterances comprising at least one lemma token included in the subset of lemma tokens;

generating, based on the natural language processing framework, a semantic relation metric indicative of semantic correspondence between the new test utterances and the test utterance sample;

determining the semantic relation metric is greater than a threshold value for only a portion of the new test utterances;

inserting, in response to the semantic relation metric being greater than the threshold value, the portion of the new test utterances in a test repository; and

testing the virtual agent based on the portion of the new test utterances.

9. The method of claim 8, wherein testing the virtual agent based on the portion of the new test utterances further comprises:

receiving a context identifier corresponding to the virtual agent, the context identifier indicative of a natural language context of the virtual agent;

compiling computer-generated test utterances associated with the context identifier; and

communicating the computer-generated test utterances to the virtual agent.

10. The method of claim 8, wherein testing the virtual agent based on the portion of the new test utterances further comprises:

communicating the portion of the new test utterances to the virtual agent, the portion of the new test utterances associated with an expected test result;

receiving, from the virtual agent, a test result responsive to a respective new test utterance in the portion of the new test utterances;

determining that the virtual agent failed a test based on a comparison of the test result with the expected test result; and

retraining a trained corpus of the virtual agent based on the portion of the new test utterances.

11. The method of claim 8, wherein obtaining the contextual corpus comprising the corpus tokens further comprises:

identifying a context identifier associated with an identifier of the virtual agent; and

identifying, in a corpus repository, the contextual corpus based in the context identifier, wherein the contextual corpus is associated with the contextual corpus in the corpus repository.

12. The method of claim 8, further comprising:

discarding, from the subset of lemma tokens, a lemma token that corresponds to corpus tokens associated with a frequency below a first threshold value;

selecting sample tokens that correspond to corpus tokens associated with a frequency above a second threshold value; and

replacing at least one of the sample tokens with at least one of the lemma tokens that were not discarded.

13. The method of claim 12, further comprising:

receiving adjustments parameters corresponding to at least one of the first threshold value or the second threshold value; and

adjusting the at least one of the first threshold value or the second threshold value based on the adjustment parameters.

14. The method of claim 8, wherein testing the virtual agent further comprises:

communicating the portion of the new test utterances to the virtual agent;

flagging, in response to an expected test result not corresponding to an actual test result, the portion of the new test utterances in the test repository;

receiving annotations configured to cause the natural language processing framework to retrain a corpus for the virtual agent;

annotating the portion of the new test utterances based on the annotations; and

communicating the annotated portion of the new test utterances to the virtual agent.

15. A non-transitory computer readable storage medium, comprising:

instructions executable by a processor, the instructions configured to cause the processor to:

obtain a test utterance sample configured to test a virtual agent;

tokenize the test utterance sample into sample tokens;

determine, based on a natural language processing framework, a synonymous set (synset) for at least one of the sample tokens, the synset comprising lemma tokens that are synonymous with the at least one of the sample tokens;

obtain a contextual corpus comprising corpus tokens;

select, from the lemma tokens, a subset of the lemma tokens that are included in the corpus tokens;

generate a plurality of new test utterances, each of the new test utterances comprising at least one of the lemma tokens included in the subset;

generate, based on the natural language processing framework, a respective semantic relation metric for each of the new test utterances, the respective semantic relation metric indicative of semantic correspondence between each of the new test utterances and the test utterance sample;

determine the semantic relation metric is greater than a predetermined similarity threshold value for similar new test utterances, the similar new test utterances being identified as similar to the test utterance sample by the predetermined similarity threshold value;

insert, in response to the semantic relation metric being greater than the predetermined similarity threshold value, the similar new test utterances in a test repository; and

communicate the similar new test utterances to the virtual agent to test the virtual agent based on the similar new test utterances.

16. The non-transitory computer readable storage medium of claim 15, wherein to determine, based on the natural language processing framework, the synset for at least one of the sample tokens, the instructions further cause the processor to:

tag each of the sample tokens with a corresponding part of speech (POS) tag;

identify at least one of the sample tokens tagged with the corresponding POS tag that is indicative of a stop word;

filter the sample tokens by removing the at least one of the sample tokens tagged with the corresponding POS tag indicative of the stop word; and

determine, based on the natural language processing framework, the synset based on at least one remaining sample token included in the filtered sample tokens.

17. The non-transitory computer readable storage medium of claim 15, wherein to select, from the lemma tokens, the subset of the lemma tokens that are included in the corpus tokens, the instructions further cause the processor to:

generate a frequency distribution of the corpus tokens;

determine, based on the frequency distribution, respective frequency values of the lemma tokens in the subset of the lemma tokens; and

select, from the subset of the lemma tokens, a highest ranked lemma token, the highest ranked lemma token being associated with a maximum one of the respective frequency values.

18. The non-transitory computer readable storage medium of claim 15, wherein to generate the plurality of new test utterances, the instructions further cause the processor to:

generate a frequency distribution of the corpus tokens;

identify at least one of the sample tokens that corresponds to at least one of the corpus tokens;

select, from the at least one of the sample tokens, a sample token that corresponds to a corpus token that is associated with a frequency value greater than a threshold value; and

replace the selected sample token with a lemma token from the subset of the lemma tokens.

19. The non-transitory computer readable storage medium of claim 15, wherein the instructions further cause the processor to:

receive a test result in response to communicating the similar new test utterances to the virtual agent;

determine that the virtual agent failed a test based on a comparison of the test result with an expected test result associated with the similar new test utterances; and

retrain a trained corpus of the virtual agent based on the similar new test utterances.

20. The non-transitory computer readable storage medium of claim 15, wherein the instructions further cause the processor to:

receive a context identifier associated with the contextual corpus;

select computer-generated test utterances based on the contextual corpus;

communicate the computer-generated test utterances to the virtual agent;

receive a test result from the virtual agent;

compare the test result with an expected test result; and

flag, in response to the received test result not corresponding to the expected test result, the computer-generated test utterances in the test repository.