
Algorithm 2 Query CTI SyscallSynonymBase

Input TargetSent /* a sentence segment in CTI document which contains STobj */
STobj /* a STobj represent the word mapping a STregex from ASG */
Output MatchedSyscalls [] /* the descriptive verb in sentence may match 0 or more system calls (1-D array) */

/* Get the descriptive verb of ST_{obj}, return if not found */

nlp ← denpency_parse(*TargetSent*, *STobj*)

verb ← find_subject_verb(*nlp*, *STobj*)

if *verb* is None **then**

return empty list []

end if

/* If the descriptive verb is in SynonymBase, get the correspondence system calls base on the verb and object type */

Initialize *SynonymList* [] ← load_synonym_base()

entityType ← classify_object_entity(*STobj*)

if *verb* ∈ query_all_verbs(*SynonymList*) **then**

return query_mapped_syscall(*SynonymList*, *verb*, *entityType*)

end if

/* If the descriptive verb is not included, get a alternative verb from SynonymBase with most similar word embedding */

/* Continue the correspondence syscall lookup process same as above */

wv ← bert_embedding(*TargetSent*, *verb*)

alternativeVerb ← query_nearest_verb(*SynonymList*, *wv*, *entityType*)

return query_mapped_syscall(*SynonymList*, *altVerb*, *entityType*)
