From Words to Action: Creating a General Narrative Planning Domain from VerbNet

Siqi Cheng and P@trik Haslum

Australian National University
firstname.lastname@anu.edu.au

1 Introduction

Research on automated story generation in AI has long had a connection with planning. Stories, like plans, are ordered sets of causally related events or actions, that lead to the story's conclusion. Planning-based methods have been used in both very early (e.g., [1]) and more recent (e.g., [2, 3]) approaches to story generation, and now incorporate many aspects of narration, such as character beliefs [4, 5] and conflict [6, 7], and audience reactions such as surprise [8] and suspense [9, 10].

The greatest obstacle, however, to successful planning-based story creation is the scale and complexity of the required planning domain, which must include all events that can take place, and all their conditions and consequences that are relevant, in some to-be-written story. The largest narrative planning domain model, to our knowledge, is that by Porteous et al. [3], with around 100 actions. This domain, however, still only models a fairly narrow story genre (hospital-based drama), and relies strongly on the tropes of this genre. It would not be usable in creating stories for other settings or genres.

Manually creating a general story domain model is a daunting task. This is why recently there is growing interest in attempting to learn domain models from stories and other texts (e.g., [11–14]).

What if there already existed a manually crafted, carefully quality-controlled, resource of formalised action knowledge, covering a broad set of actions and events? VerbNet [15, 16] is such a resource. It is a lexicon of around 4,500 English verbs, organised into over 600 distinct classes, annotated with both syntactic and semantic information. In this paper, we investigate the extent to which VerbNet's semantic annotations enable translating these verb classes into actions of a planning domain. The result is a domain with 432 distinct action schemas, 55% of which have at least one effect that unifies with a precondition of some action.[1] While this is substantial in scope, the domain also has significant limitations, some due to our construction and some due to VerbNet's annotations. We believe that combining VerbNet with linked lexical resources and domain learning methods can address some of those limitations.

[1] Code and the resulting domain are available at https://github.com/chengsiqi-37/verbnet-2-actionmodel.

2 VerbNet

VerbNet [15, 16] is a lexicon of English verbs annotated with both syntactic and semantic information. It is organised into *classes*, which collect sets of verbs that take the same complements (arguments), allow the same syntactic frames, and have a similar meaning. The same word (verb) can appear in several classes, reflecting that words can often have more than one meaning. For example, the class accompany-51.7, which describes an entity moving along with another, contains the verbs "accompany", "conduct", "escort", "guide", "lead", "shepherd" and more, but the verb "lead" also appears in classes result-27.2 (meaning "lead to" or "result in"), supervision-95.2.2-1, and more.

The semantic annotation of a VerbNet class has two parts: The first is a set of *thematic roles*, which describe participants in the event. For example, roles in accompany-51.7 are an *Agent* (the one who accompanies) and a *Theme* (the one who is being accompanied), as well as optional roles providing specific details, such as *Initial_Location*, *Trajectory* and *Destination*. Thematic roles form a taxonomy that defines the type of argument that can fill them, and in a verb class can also be specialised with *selectional restrictions*, which are properties that the argument must have, or not have. For example, in accompany-51.7, both *Agent* and *Theme* have the restriction *+animate*.

A *syntactic frame* describes a particular syntactic structure that verbs of the class can participate in, and defines how syntactic constituents (subject, object, etc) map to the thematic roles. Not all thematic roles are necessarily present in every frame. For example, in the basic transitive frame of accompany-51.7, "*Agent* Verb *Theme*", all the optional spatial roles are unspecified.

Associated with each frame is a set of *semantic predicates*. The semantic representation has been significantly revised in the most recent release of VerbNet [17, 18]. Viewed through a planning lens, we distinguish three kinds of predicates:

– Predicates that express a property of or relation between participants in the event at a time point or over a time interval. We call these *fluent* predicates. The time points or intervals are called *subevents* by Brown et al. [18].
– Predicates that express a relation between roles without reference to a specific time (e.g., part_of), implying that the relation is unchanging, at least in the context of the event. We call these *static* predicates.
– Predicates that express a property of or relation between the subevents (time points or intervals) involved in the event. The majority of these express a temporal relation (such as start or co-temporal), but they also include (cause e,e'), denoting that subevent e is the cause of e', and (irrealis e), denoting that subevent e is not real (i.e., an expected, hypothetical, or counterfactual subevent). Nevertheless, we call these *temporal* predicates.

The (non-temporal) arguments of fluent and static predicates in a frame are a mix of explicit and implicit thematic roles, and verb-specific constants. It is not always clear what the intended type of an argument is.

According to Brown et al. [17], subevents are implicitly ordered by their numbering. We interpret this order as non-strict. For example, the semantic representation of the accompany-51.7 event is depicted in Figure 1.

(has_location e_1, *Agent*, *Initial_Location*)

(has_location e_2, *Theme*, *Initial_Location*)

(motion $\ddot{e}_3$, *Agent*, *Trajectory*), $\neg$(has_location $\ddot{e}_3$, *Agent*, *Initial_Location*)

(motion $\ddot{e}_4$, *Theme*, *Trajectory*), $\neg$(has_location $\ddot{e}_4$, *Theme*, *Initial_Location*)

(co-temporal $\ddot{e}_3$, $\ddot{e}_4$)

(has_location e_5, *Agent*, *Destination*)

(has_location e_6, *Theme*, *Destination*)

Fig. 1. Semantic representation of the accompany-51.7 verb class. The notation $\ddot{e}$ indicates that the subevent is a time interval or process. Note, however, that a subevent without this annotation is not necessarily instantaneous, and can also extend in time.

Verb classes can be in a super-/subclass relationship, in which the subclass specialises the superclass by having additional selectional restrictions on some roles, by admitting additional frames, or by extending some frame's semantic description with additional predicates. The class hierarchy (VerbNet version 3.4) extends to at most five levels of depth, while there are 329 top-level classes.

3 Planning Domain Creation

Among the fluent predicates used in VerbNet's semantic representations, some clearly describe an aspect of the state of an event participant (e.g., has_location or alive), while some describe an activity (e.g., motion or transfer). Since a classical planning model describes only the state changes that result from an action, not what takes place during the action, we include only the former in the action models. We manually partition the fluent predicates into *state* and *activity* predicates, and use the first occurrence of an activity predicate in a frame's semantic representation as the boundary between its pre- and postcondition: state predicates with subevent labels preceding that of the first activity predicate are interpreted as preconditions, while those at the same or subsequent subevents are interpreted as postconditions. Certain cases require special handling:

- *Frames without an activity predicate*: Some frames (e.g., adopt-93.1) do not contain any activity predicate. In those cases, we treat all predicates as postconditions (except as potentially modified in the next case).

- *Contradictory effects*: In some cases – both that have an activity predicate and that do not – we end up with contradictory literals in the postconditions, which clearly signals a change of state. In these cases, we assign the predicate with the earlier subevent label to the precondition, and the contrary one to the postcondition. Importantly, we found no instances of frames that already have preconditions and that also have contradictory effects.
- *Static predicates*, such as part_of or cost, are not associated with a subevent. We treat these as preconditions.

The total number of (fluent and static) state predicates used in the planning domain is 85.

Action Schemas We initially create one action schema for each frame of each of verb class. The schema's parameters include all thematic roles in the frame's semantic representation, both explicit and implicit. We classify each predicate argument as a thematic role if it is explicitly marked with the *ThemRole* type or it is found in the thematic role hierarchy of VerbNet's documentation[2]; all other arguments are considered constants. This approach identified 46 distinct thematic roles occurring in semantic annotations.

VerbNet version 3.4 contains a total of 602 classes (including subclasses)[3], comprising 1591 frames. After removing actions with no effects, we obtain 1166 action schemas. 87 verb classes yield no action schema, due to having no state predicate in their effects.

Because frames are a syntactic concept, the difference between frames of the same verb class is often just whether a role is explicitly mentioned or not, or even a syntactic variation without semantic difference (e.g. "I teach logic to students" vs. "I teach students logic"). Thus, different frames of a verb class often have identical semantic representations, resulting in identical action schemas.

To obtain unique action schemas, we merge actions derived from different frames belonging to the same verb class or its subclasses that have equal preconditions and effects. Since thematic roles are sometimes narrow, we generalise action schemas by merging frames with semantically similar roles. For instance, the predicates (together *Patient*, *Patient*) and (together *Patient*, *Co-Patient*) are treated as equivalent, as both *Patient* and *Co-Patient* are subtypes of the broader *Undergoer* role. Subevent labels, and for each other argument whether they are explicit or implicit, is ignored. This yields a total of 432 unique action schemas across all verb classes. The class with the highest number of distinct action schemas is break, with a total of 6.

Connectivity To evaluate the internal connectivity of the domain, we analyse how frequently predicates appear in pre- and postconditions. Only 48 predicates

[2] https://uvi.colorado.edu/references_page
[3] A labeling inconsistency was identified in the dataset, where a subclass of bill-54.5 was incorrectly labeled with the same class ID as its parent class, bill-54.5. To ensure that the subclass is correctly included in the analysis, we relabeled it as bill-54.5-1 without modifying any additional data.

occur in both precondition and effects. However, 238 of the 432 action schemas (55%) have at least one effect that unifies with a precondition of some action schema, implying that these actions can be part of a causally connected plan. Conversely, 187 action schemas (43%) have at least one precondition that unifies with an effect.

4 Limitations and Potential

The planning domain we have created from VerbNet is a first try, with several limitations. The classical planning restriction to state predicates only omits sometimes significant information. For example, the semantic difference between the frames "*Patient* Verb" (e.g., "the noise increased") and "*Causer* Verb *Patient*" (e.g., "they increased the noise") of caused_calibratable_cos-45.6.2-1 is only that in the latter, the change of state in *Patient* is caused by an action done by the *Causer*; since we classify do and cause as activity predictes, this distinction is lost. Moreover, in the narrative planning context, transient action effects can indirectly alter the state, for instance the mental state of a character that witnesses the action.

Each verb class models a specific verb sense, at a certain level of abstraction. For example, accompany-51.7 does not distinguish if the *Theme* is a willing or forced participant in the action. Thus, despite its impressive scope, VerbNet covers only a fraction of potential actions. Furthermore, to fully interpret the semantic annotations requires additional knowledge. For instance, verb classes hit-18.1 and touch-20 both have the effect of changing (contact *Agent*, $\cdot$) from false to true, with the difference that in hit-18.1 the manner of contact is "forceful". The potential consequences of this manner, e.g., that one or both objects may break or hurt, or cause the object hit to move, however, remain implicit.

On its own, VerbNet is not sufficient to create a general narrative planning domain. However, it may yet be of great use in combination with other methods. We plan to investigate its integration with domain learning methods, for example as a source of distant supervision and as a base on which to learn additional facts. A potentially valuable feature of VerbNet is that in addition to its annotations, it also records links with other lexical resources, such as WordNet and FrameNet, which may be used to identify candidate novel verb classes.

References

1. Lebowitz, M.: Story-telling as planning and learning. In: Proc. 10th International Joint Conference on Artificial Intelligence (IJCAI). (1983)
2. Riedl, M., Young, R.: Narrative planning: Balancing plot and character. Journal of AI Research **39** (2010) 217–268
3. Porteous, J., Charles, F., Cavazza, M.: NetworkING: using character relationships for interactive narrative generation. In: International conference on Autonomous Agents and Multi-Agent Systems (AAMAS). (2013) 595–602

4. Teutenberg, J., Porteous, J.: Incorporating global and local knowledge in intentional narrative planning. In: 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS). (2015)
5. Christensen, M., Nelson, J.M., Cardona-Rivera, R.E.: Using domain compilation to add belief to narrative planners. In: AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE). (2020) 38–44
6. Ware, S.G., Young, R.M.: Glaive: A state-space narrative planner supporting intentionality and conflict. In: AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE). (2014)
7. Porteous, J., Lindsay, A.: Protagonist vs antagonist PROVANT: Narrative generation as counter planning. In: 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS). (2019) 1069–1077
8. Bae, B.C., Young, R.M.: A computational model of narrative generation for surprise arousal. IEEE Transactions on Computational Intelligence and AI in Games **6**(2) (2014) 131–143
9. Cheong, Y.G., Young, R.M.: Suspenser: A story generation system for suspense. IEEE Transactions on Computational Intelligence and AI in Games **7**(1) (2015) 39–52
10. Winer, D.R., Young, R.M.: Discourse-driven narrative generation with bipartite planning. In: 9th International Natural Language Generation conference. (2016) 11–20
11. Hayton, T., Porteous, J., Ferreira, J., Lindsay, A., Read, J.: StoryFramer: From input stories to output planning models. In: ICAPS Workshop on Knowledge Engineering for Planning and Scheduling (KEPS). (2017)
12. Huo, Y., Tang, J., Pan, Y., Zeng, Y., Cao, L.: Learning a planning domain model from natural language process manuals. IEEE Access **8** (2020) 143219–143232
13. Miglani, S., Yorke-Smith, N.: Nltopddl: One-shot learning of pddl models from natural language process manuals. In: ICAPS Workshop on Knowledge Engineering for Planning and Scheduling (KEPS). (2020)
14. Li, R., Cui, L., Lin, S., Haslum, P.: NaRuto: Automatically acquiring planning models from narrative texts. In: 38th AAAI Conference. (2024) 20194–20202
15. https://uvi.colorado.edu/
16. Kipper Schuler, K.: VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon. PhD thesis, University of Pennsylvania (2005)
17. Brown, S.W., Pustejovsky, J., Zaenen, A., Palmer, M.: Integrating generative lexicon event structures into VerbNet. In: Proc. International Conference on Language Resources and Evaluation (LREC). (2018) 56–61
18. Brown, S.W., Bonn, J., Gung, J., Zaenen, A., Pustejovsky, J., Palmer, M.: VerbNet representations: Subevent semantics for transfer verbs. In: Proc. 1st International Workshop on Designing Meaning Representations. (2019) 154–163