Computational Linguistics and Chinese Language Processing
Vol. 19, No. 2, June 2014, pp. 33-52

Resolving the Representational Problems of Polarity and Interaction between Process and State Verbs

Shu-Ling Huang[*], Yu-Ming Hsieh[*+], Su-Chu Lin[*], and Keh-Jiann Chen[*]

Abstract

Event classification is one of the crucial tasks in lexical semantic representation. Traditionally, researchers have regarded process and state as two top-level events and discriminated between them by semantic and syntactic characteristics. In this paper, we add cause-result relativity as an auxiliary criterion to discriminate between process and state by structuring about 40,000 Chinese verbs to the two correspondent event hierarchies in E-HowNet. All verbs are classified according to their semantic similarity with the corresponding conceptual types of ontology. As a result, we discover deficiencies of the dichotomy approach and point out that any discrete event classification system is insufficient to make a clear-cut classification for synonyms with slightly different semantic focuses. We then propose a solution to remedy the deficiencies of the dichotomy approach. For the process or state type mismatched verbs, their inherited semantic properties will be adjusted according to their PoS and semantic expressions to preserve their true semantic and syntactic information. Furthermore, cause-result relations will be linked between corresponding processes and states to bridge the gaps of the dichotomy approach.

Keywords: Event Classification, Process and State, Lexical Representation, Cause-result Relativity between Verbs.

1. Introduction

Clarifying the nature of verb classes is a crucial issue in lexical semantic research, being of great interest to both theoretical and computational linguistics. Many classification and representation theories have been presented already, including the widely cited theories

[*] Institute of Information science, Academia Sinica, Taipei, Taiwan
[+] Department of Computer Science, National Tsing-Hua University, Taiwan
E-mail: {josieh,morris,jess,kchen}@iis.sinica.edu.tw

proposed by Vendler (1967), Dowty (1979), Bach (1986), Parsons (1990), Levin (1993), Pustejovsky (1995), and Rosen (2003). Additionally, several online verb classification systems, such as WordNet (Fellbaum, 1998), VerbNet (Kipper-Schuler, 2006), FrameNet (Fillmore *et al*., 2003), and Levin's verb classification also are available. Each approach views events from a different perspective, and each approach clarifies a different part of the overall problem of understanding the linguistic representation of events. Overall, they can be divided into two main schools, one is semantic classification, such as Vendler's approach, and the other is syntactic classification, such as Levin's approach.

Since different event classifications pinpoint the basic features of events that need to be represented, we need to clarify the goal we want to achieve before adopting or proposing an event classification. In this paper, we aim to achieve a better lexical semantic representation framework for E-HowNet (Chen *et al*., 2003), and we adopt the typologies of process and state as the two top-level event types. Since verbs may express different aspects or viewpoints of conceptual events, however, it is difficult to make a clear-cut difference between process and state verbs in some cases. Verb-result compounds, such as 購妥 gou-tuo 'to complete procurement,' are obvious examples of being either pure process or state. Furthermore, semantic interactions of the verbs also need to be clarified. Consider, for example, the synonymous words (strictly speaking, near synonyms and hyponyms) of 記得 ji-de 'remember' in Mandarin Chinese: (a) 想起 xiang-qi 'call to mind,' 記取 ji-qu 'keep in mind,' 背起來 bei-qi-lai 'memorize,' (b) 念念不忘 nian-nian-bu-wang 'memorable,' and 刻骨銘心 ke-gu-ming-xin 'be remembered with deep gratitude'. Although these words are near synonyms, their senses shift slightly according to different semantic focuses, often resulting in different grammatical behavior. If we classify Group (a) as a process type, and Group (b) as a state type by their fine-grained semantic focuses, we may lose the important information that they are actually near synonyms and have the core sense of 記得 ji-de 'remember'. Therefore, in order to design a better semantic and syntactic representational framework for verbs, we try to clarify the polarity and interaction between process and state.

The remainder of this article is organized as follows. In the next section, we begin with a review of past research. Section 3 clarifies the polarity between process and state before addressing difficulties of the dichotomy approach. In Section 4, we describe the interaction between process and state, propose solutions to overcome the difficulties mentioned in the previous section, and discuss other event relations that should be represented in analogy with process state dichotomy. Finally, we conclude our findings and possible future research in Section 5.

2. Background

Over 2300 years ago, Aristotle (in Jonathan Barnes eds., 1984) proposed the first event-based classification of verbs. His main insight was the distinction between states and events (called 'processes' in this paper). Since the late 1960s, a large number of event classifications, variously based on temporal criteria (such as tense, aspect, time point, and time interval), syntactic behavior (such as transitivity, object case, and event structure), or event arguments (such as thematic role mapping, agent type, and verb valence) have been suggested and have aroused heated discussion. These representations can be roughly divided into the two main schools of semantic classification and syntactic classification. In the following discussion, we take Vendler and Levin as representatives for the two schools and we find that both schools treat process and state as two clearly different event types.

2.1 Vendler's Classification

Vendler's classification (1967) is the most influential and representative system in terms of the semantic classification approach. He classified verbs into four categories "to describe the most common time schemata implied by the use of English verbs" (pp. 98-99). The four categories are given in (1).

(1) a. *States*: non-actions that hold for some period of time but lack continuous tenses.

b. *Activities*: events that go on for a time, but do not necessarily terminate at any given point.

c. *Accomplishments*: events that proceed toward a logically necessary terminus.

d. *Achievements*: events that occur at a single moment; therefore, they lack continuous (progressive) tenses.

Distinctly, states denote a non-action condition and are irrelevant to temporal properties, while the other three denote an event process or a time point in an event process. Vendler's successors, such as Verkuyl (1993), Carlson (1981), Moens (1987), and Hoeksema (1983), extended this discussion without changing Vendler's basic framework. According to Rosen (2003), the successors all pointed out that state and process are two major event types. Ter Meulen (1983; 1995) thus suggested a redefinition of Vendler's classes. She defined states as having no internal structure or change, while events, i.e., the processes dealt with in our paper and composing Vendler's other three event types, are defined on the basis of their parts.

2.2 Levin's Classification

Levin (1993) believes that identifying verbs with similar syntactic behavior provides an effective means of distinguishing semantically coherent verb classes. She proposed a coarse-grained classification for verbs based on two observations: the first is that many result verbs lexicalize results that are conventionally associated with particular manners, and vice-versa, many manner verbs lexicalize manners that are conventionally associated with particular results. The examples she gave are listed in (2):

(2) The pervasiveness of the dichotomy (Levin, 2011)

	Manner verbs	**vs.**	**Result verbs**
Verbs of damaging:	*hit*	vs.	*break*
Verbs of putting—2-dim	*smear*	vs.	*cover*
Verbs of putting—3-dim	*pour*	vs.	*fill*
Verbs of removal	*shovel*	vs.	*empty*
Verbs of combining	*shake*	vs.	*combine*
Verbs of killing	*stab*	vs.	*kill*

Levin argued the origin of the dichotomy arises from a lexicalization constraint that restricts the manner and result meaning components to fit in a complementary distribution: a verb lexicalizes only one type and those components of a verb's meaning are specified and entailed in all uses of the verb, regardless of context. Further, not only do manner and result verbs differ systematically in meaning, but they differ in their argument realization options (Rappaport & Levin, 1998; 2005). For example, result verbs show a causative alternation, but manner verbs do not, as shown in Example (3); and, manner verbs show considerably more and different argument realization options than result verbs (Rappaport & Levin, 1998), such as those described in (4).

(3) a. Kim broke the window./The window broke.

b. Kim wiped the window./*The window wiped.

(4) a. Terry wiped. (activity)
b. Terry wiped the table. (activity)
c. Terry wiped the crumbs off the table. (removing)
d. Terry wiped the crumbs into the sink. (putting)
e. Terry wiped the slate clean. (change of state)
f. Terry wiped the crumbs into a pile. (creation)

Levin's manner verb and result verb dichotomy characterizes semantic and syntactic interactions between verbs. Specifically, this syntactic dichotomy is caused by the semantic characteristics of the language. We consider a similar semantic relation of cause-result between process verbs and state verbs to show the dichotomy and interactions between them. In fact, Levin's result verbs are verb-result compounds in Chinese, such as the process verb 打破 da-po 'break' in our classification. We regard results of processes to be result states, such as 破裂 po-lie 'broken'. Hence, the aforementioned verb pairs, such as *stab* and *kill* in (2), are both process verbs. By our notion of process and state dichotomy, wounded and die are result states of *stab* and *kill*, respectively.

2.3 E-HowNet's Classification

E-HowNet (Chen *et al*., 2005) is a frame-based entity-relation model that constructs events, objects, and relations in a hierarchically-structured ontology. By following the conventional event classification theories, verbs are partitioned into process and state first, which is a higher priority dichotomous classification criterion than the syntactic classification in E-HowNet, since E-HowNet primarily is a semantic classification system. Furthermore, semantic classification is more intuitive and more in line with the general view of the real world. Based on this criterion, the top-level E-HowNet ontology is established, as depicted in Figure 1, and a snapshot of E-HowNet is given in Appendix A.

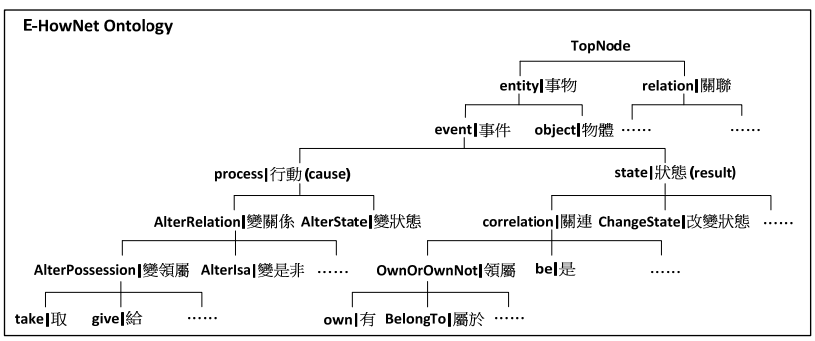


Figure 1. The Architecture of E-HowNet

3. The Polarity and Interaction between Process and State

Process and state have long been treated as two top classes of events. Semantically, their distinctions are evident and intuitive, such as the difference between the process verb 取悅 qu-yue ‘please’ and the state verb 喜悅 xi-yue ‘joyful’. With respect to syntax, process and state verbs also have their own individual characteristics; for example, 取悅 qu-yue ‘please’ must have a patient object but 喜悅 xi-yue ‘joyful’ does not. Differentiating them is considered obvious in theoretical and practical linguistic research areas. Nevertheless, from the perspective of a fine-grained lexical analysis, researchers have also found that it is difficult to make clear-cut differences between process and state. Take the following as examples. The state verb 生氣 sheng-qi ‘angry’ may accept an object goal in Mandarin and is difficult to differentiate from the process verb 發脾氣 fa-pi-qi ‘get angry’ in semantics. In this paper, we do not aim to strictly partition 生氣 sheng-qi ‘angry’ and 發脾氣 fa-pi-qi ‘get angry’ into state and process type. Instead, our objective is to discriminate processes from states with an emphasis on why we encounter difficulties of discriminating them and what better representations may preserve as much semantic and syntactic information as possible. For example, the verb 遇害 yu-hai ‘be murdered’ can be either classified as a process of *kill* or a state of *die*, with neither classification being absolute. A better solution might be that, even if the verb is misclassified into either type, we can still recognize that the experiencer of 遇害 yu-hai ‘be murdered’ is killed and dead. In this section, we emphasize the general distinction between process and state. Then, in the next section, we introduce several approaches we adopted upon encountering difficulties of process-state dichotomy.

The differentiating characteristics between process and state verbs, other than semantic differences, are not obvious. Summarizing the previously mentioned theories in Section 2, the polarities between process and state can be generalized as follows.

(5) The polarities and interactions between process and state

Processes: cause of states, dynamism (*i.e.*, relevant to temporal properties), object domination

States: result of processes, stasis (*i.e.*, irrelevant to temporal properties), object modification

The polarity of dynamism and stasis is a semantic-based distinction, whereas the domination of objects or their modification is a syntax-based distinction. They are both common but coarse-grained event classification criteria, and most verbs can be distinguished by these coarse-grained classification criteria. Nevertheless, some verbs, like 發脾氣 fa-pi-qi ‘get angry’ and 遇害 yu-hai ‘be murdered,’ are not classified easily. In our study, we propose

an interaction between cause and result as an auxiliary criterion, which asserts that *processes* are the cause of states and they denote an event process or a time point on an event process. On the other hand, *states* are the result of processes and they denote a non-action condition and are irrelevant to temporal properties, *i.e*., they have no internal structure or change. Although it would appear that cause-result is a natural differentiation criterion between processes and states, it may not be a one-to-one relation and some verb types may not have obvious cause-result counterparts. For instance, the concept of causative process {earn|賺} may achieve several resultant states, such as {obtain|得到} and {rich|富}, although the process of {swim|游} does not have an obvious result state. Nonetheless, if we can use the characteristics of (5) to differentiate all verbs into process and state types, it may help us achieve the first step towards a lexical semantic classification for verbs. We then use semantic expressions, part-of-speech (PoS) features,[1] and relational links, such as cause-result relationship between process types and state types, to make a better lexical semantic representation. Regarding the verb type classification, the following questions may be raised. Is the process-state dichotomy approach feasible? How are the verbs denoting complex event structures, such as verb-result compounds, classified? Is it true that all states have causing processes and all processes have resulting states? The following observations will provide the answers to these questions.

3.1 Observations and Difficulties of the Process-State Dichotomy in E-HowNet

In order to develop the lexical semantic representation system E-HowNet, we classified all Chinese verbs into a process and state type hierarchy, as illustrated in Figure 1. We use the characteristics (5) of dynamism and stasis as semantic-based distinctions, the domination and modification of objects as syntax-based supporting criteria, and the cause-result relation as a complementary criterion to distinguish process from state. It is interesting that, with the exception of general acts, almost all top-level Chinese verb types, whether of process or state types, necessarily have their cause-result counterpart. Nevertheless, for the fine-grained lower level types or lexical level verbs, there are three different cases of lexical realizations of cause-result dichotomy, which are listed in the following.

Case 1: Process types have result states and *vice-versa*. An example of cause-result mapping between process and state is given in (6).

[1] For simplicity, in this paper, we only tag the top-level PoSs, *i.e*. active PoS and stative PoS, which are adopted from the classification of CKIP group (1993).

(6) Causative process type {brighten|使亮}: *e.g.*, 磨光 mo-guang ‘burnish’, 擦亮 ca-liang ‘polish’ ←→
Resultant state type {bright|明}: *e.g.*, 水亮 shui-liang ‘bright as water’, 光燦 guang-can ‘shining’

For this case, the process and state are two different types and can be differentiated by the fundamental differences between dynamic and static types or by the cause-result relation. Nevertheless, lexemes may shift their senses due to different compounding, resulting in a classification dilemma of semantic similarity first or dichotomy of process and state first. As was mentioned in the above example, the causative process type {kill|殺害}, *e.g.*, 弔死 diao-si ‘hang by the neck,’ has a resultant state type {die|死}, *e.g.*, 往生 wang-sheng ‘pass away’. Then, how about the result-state verb 遇害 yu-hai ‘be murdered’? Should we classify 遇害 yu-hai ‘be murdered’ as a process type {kill|殺害} or as a state type {die|死}? The verb 遇害 yu-hai ‘be murdered’ seems to be the resultant state {die|死} in terms of stativity, but from the perspective of a semantic focus, it is more akin to a causative process {kill|殺害}. This classification difficulty always occurs when we analyze verbs denoting different aspect situations, such as passive or achieved aspects. As a result, near synonyms of the same event type could be separated for denoting different aspectual situations.

In terms of the E-HowNet ontology, the cause-result matching between processes and states almost reaches 100% respecting hypernymy concepts exemplified by corresponding lexical pairs, as shown in Figure 2. Nevertheless, at the hyponym or lexical level, we found that the correspondent rate was not as high as in top-level concepts. This results in Case 2 below.

Case 2: Process types do not have nodes of result states nor do state types have nodes of causing processes in the E-HowNet ontology, which means the result states or causal processes are either vague or they are not lexicalized common concepts. (7), (8) are typical examples.

(7) The causative process type {punish|處罰}, such as 行刑 xing-xing ‘execute’ or 處決 chu-jue ‘put to death,’ have corresponding aspectual resultant states, such as 受刑 shou-xing ‘be put to torture’ and 伏法 fu-fa ‘be executed,’ but no lexicalized concept in common to denote *being punished* or *being tortured* in Chinese. Therefore, there is no proper node of state type to which the above two stative verbs belong in E-HowNet.

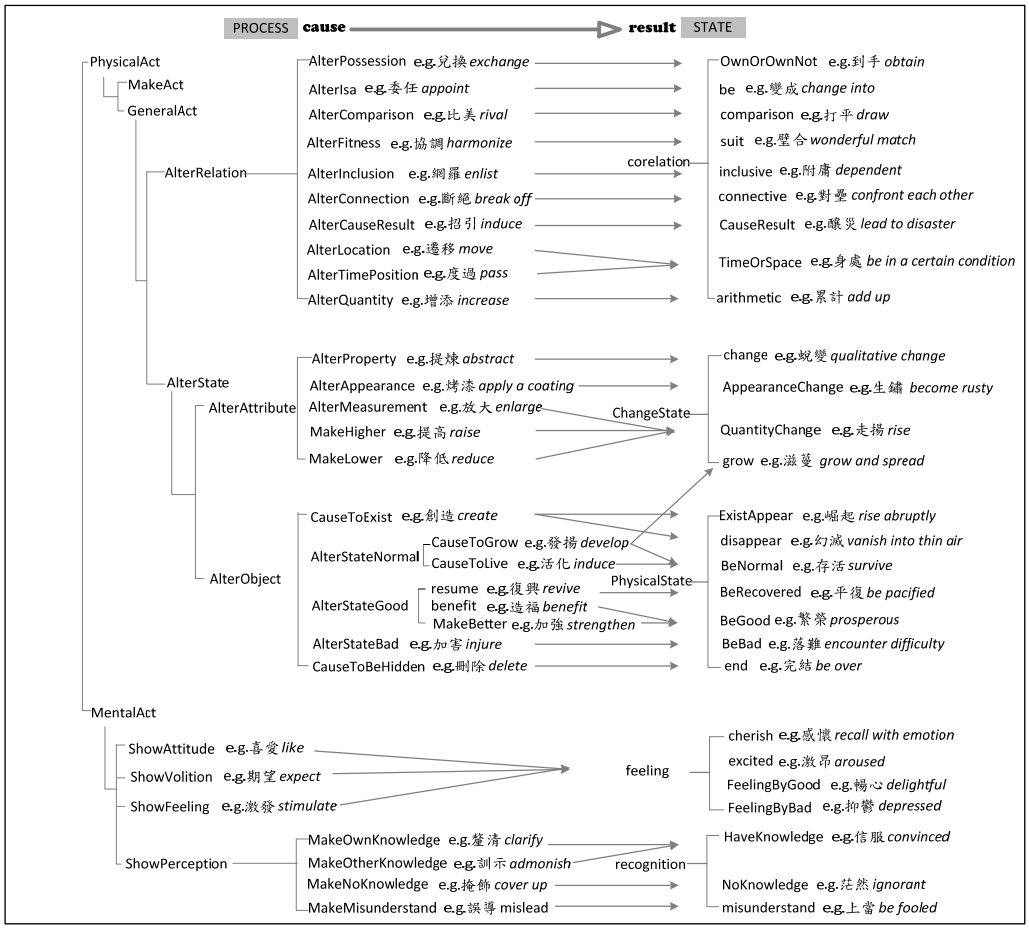


Figure 2. The Matching between Processes and Result States in E-HowNet

(8) There is no lexicalized concept in common to denote causative processes, such as 板起(臉) ban-qu-(lian) 'put on a stern expression' and 正色 zheng-se 'with a stern countenance' in Chinese, which are the cause of the resultant state type {austere|冷峻}, *e.g.*, 凝重 ning-zhong 'serious,' 不苟言笑 bu-gou-yan-xiao 'serious in speech and manner'. That is, there is no proper node of the process type to which the above two process verbs 板起臉 ban-qi-lian 'put on a stern expression' and 正色 zheng-se 'with a stern countenance' belong.

For lexemes of Case 2, the characteristics of process and state of (5) can still differentiate the lexemes on the process and state types, but there are no actual corresponding conceptual nodes in the ontology. This means that some stative verbs must be attached to the process type node and some process verbs should be attached to stative type nodes in the ontology for the

sake of keeping reasonable semantic consistency.

Case 3: Some processes and respective states co-exist concurrently and are not in the cause-result temporal sequence. We call such a concurrent process and state a dual process-state. There are 22 dual process-state type primitives in the E-HowNet ontology (refer to Appendix B), with Example (9) describing one of them.

> (9) The dual process-state {living|生活} includes: (a) 求生 qiu-sheng ‘seek to survive,’ 度日 du-ri ‘subsist,’ and (b) 生存 sheng-cun ‘exist,’ 在世 zai-shi ‘be living,’ 一息尚存 yi-xi-shang-cun ‘be still alive’. The semantic focus of (a) indicates a process of *making a living* or *to live*, while (b) indicates the state of *being alive* or *be living*. The two types of process and state coexist and they do not have cause-result relation.

For the dual process-state type, we encounter the similar dilemma of the previous two cases. If we choose the bipartite process and state approach, near synonyms will belong to two nodes far apart in the ontology. If we adopt the approach of a unified conceptual node for each dual process-state type, the result will be the same problem as in Case 2, *i.e.*, stative verbs and process verbs are of the same type.

Furthermore, in Mandarin Chinese we have many verb-result compounds (VR), such as 累病 lei-bing ‘sick from overwork,’ 驚退 jing-tui ‘frighten off,’ and 購妥 gou-tuo ‘to complete procurement’. Since the causative process and resultant state are contained in the same verb, how should we classify them?

4. Knowledge Representation for Process and State Verbs

The difficulties of the dichotomous approach are caused by the semantic interaction between state and process. We thus propose the classification criterion (5) and a representational scheme according to the above observations, and we try to solve the difficulties without changing the framework of the dichotomy structure. The idea is that all verbs are classified into the most similar conceptual types, according to their respective sense. The process or state type mismatched verbs will have their types adjusted by their PoS or semantic expressions. Such an approach is functional, like using the feature of ‘*don’t fly*’ to adjust the flying property for penguins as bird type and still maintaining the inherent properties. Furthermore, cause-result relations will be established between corresponding processes and states to bridge the gaps of the dichotomy approach.

4.1 Lexical Semantic Representation for Process and Stative Verbs

For the Case 1 verbs, every process has a corresponding result state, and every state has a corresponding causal processes. For synonymous verbs with a process and state dichotomy, each verb is placed under its corresponding conceptual node. In addition, the cause-result relationship will be established between corresponding process types and state types, as exemplified in Figures 2 and 4. In real implementation, there are 310 corresponding cause-result pairs established. Nevertheless, from a practical point of view, all semantic representation systems are discrete systems. Given that they use a limited number of primitive concepts to express complex concepts, the result is that some words are forced to be classified to the most similar concept node but with a mismatched major semantic type, such as 遇害 yu-hai 'be murdered' possibly being classified as the process type {kill|殺害} instead of the state type {die|死}. We will resolve such problems by following the same method for Case 2 verbs.

As shown in the observation of Case 2, some of the cause-result corresponding concepts are vague and some are not lexicalized, neither of which occurred as conceptual nodes in the ontology. As a result, for verbs whose potential hypernyms are missing, we will classify these verbs to their cause-result counterpart conceptual nodes instead. After that, we use the part-of-speech to recover the correct semantic type of state or process, as exemplified in (10).

(10) Causative process: {FondOf|喜歡} ←→ there is no corresponding resultant state

The typical examples of semantic type of {FondOf|喜歡} are 看中 kan-zhong 'take fancy to,' 喜愛 xi-ai 'love,' 酷愛 ku-ai 'ardently love,' and 熱衷 re-zhong 'be addicted to'. They are tagged with an active PoS. The verbs 癡情 chi-qing 'be infatuated' and 興致盎然 xing-zhi-ang-ran 'full of interest,' however, are stative verbs, but there is no lexicalized state primitive to place these verbs. Hence, they are classified to the most similar hypernym concept node, *i.e.*, {FondOf|喜歡}.

With part-of-speech tags, we have no problem discriminating state verbs that are attached to a process primitive. In fact, we can define state verbs in {result({process})} format; or process verbs in {cause({state})} format in order to make both semantic distinctions and link relations.

(11) 看中 kan-zhong 'take fancy to,' 喜愛 xi-ai 'love,' 酷愛 ku-ai 'ardently love,' 熱衷 re-zhong 'be addicted to' are defined as {FondOf|喜歡};

癡情 chi-qing 'be infatuated,' 興致盎然 xing-zhi-ang-ran 'full of interest' are defined as {result({FondOf|喜歡})} and have stative PoS.

Moreover, fine-grained part-of-speech tags also provide syntactic information for each verb; this solves the difficulty of Case 2 and effectively expresses fine-grained semantic and syntactic distinctions for near-synonyms.

4.2 Lexical Representation for Dual Process-State Verbs

For Case 3 dual process-state verbs, the bipartite classification for state and process no longer exists for two reasons. First, it is difficult to make a distinction between process and state for the dual types. Second, state and process are just two different viewpoints of the same events. A single dual process-state conceptual type may contain both process and stative verbs of the same event type but different viewpoints. We use part-of-speech tags to tell the difference between semantic focus and the syntactic behavior of each verb. In addition, the dual process-state type also indicates that the process and state coexist at the same event duration. For instance, both verbs 度日 du-ri ‘subsist’ and 在世 zai-shi ‘be living’ are belong to the same conceptual type of {living|生活}, but have the active PoS and stative PoS, respectively.

4.3 Lexical Semantic Representation for Verb-Result Compounds

In addition to the verbs belonging to Cases 1-3, we also wanted to address the solution for classification difficulty of VR compounds. A VR compound may be a composition of a process event followed by an event of result state, such as 打破 da-po ‘break’. The verb in (12.a) is more process-like, but the same verb in (12.b) is more state-like. It is a dilemma to classify the verb into either process or state.

(12.a) 張三打破花瓶 Zhang-san da-po hua-ping ‘Zhang San broke the vase.’

(12.b) 花瓶打破了 hua-ping da-po le ‘The vase is broken.’

Nevertheless, if its semantic expression provides sufficient information to clarify the accurate word meaning and relation between V1 and V2, as well as a suitable PoS classification, there is no difference in the event type where it was classified. Although it is controversial to recognize the semantic focus of these verbs, *i.e.*, to determine whether they are more state-like or more process-like, it is not an important issue in making a semantic and syntactic distinction in lexical representation. We built explicit links of cause-result relations between sub-events in the LESRE framework of E-HowNet (Chen *et al.*, 2013). For example, the sense of VR verb 驚退 jing-tui ‘frighten off’ is expressed as in (13). We also encoded the co-indexed arguments for all related event pairs, *i.e*. the patient of {frighten|嚇唬} is the agent of {leave|離開} in (13).

(13) 驚退 jing-tui ‘frighten off’ def:{frighten|嚇唬: patient={x}, result={leave|離開:agent={x}}}

In order to maintain fluency and legibility of the article, the PoS features and semantic expressions of all of our examples are listed in Appendix C.

4.4 Lexical Representation for Linking Semantic Related Concepts

The connection of semantic relations between concepts is almost as important as the classification of events in a hierarchical framework. Since the construction of a hierarchical taxonomy is primarily by hypernym-hyponym relations, many semantically related concepts may be far apart in the taxonomy, such as cause process and result state. Therefore, we must also take semantic connection and fine-grained lexical representation into account when classifying events into groups. There are 11 types of explicit relations in HowNet identified by Dong & Dong (2006), also adopted by E-HowNet, to link the related concepts. They are synonym, synclass, antonym, converse, hypernym, hyponym, cognate role-frame, part-to-whole, value-to-attribute, attribute-to-host, and semantic-roles-to-event. In fact, the supplement linking relations between two semantically related but hierarchically far apart concepts in E-HowNet are more than the aforementioned relations. We use E-HowNet expressions to express semantic equivalence and link the two concepts. For instance, for the related concepts of *able* and *ability*, *bad* and *good* their relations can be expressed as ability=degree({able|能}) and {bad|壞}=not({nice|良好}). In this paper, we find processes and states exist with a cause-result relation that can be expressed in a function form as result({act|行動})={state|狀態} or cause({state|狀態})={act|行動}, such as result({CauseToAppear|顯現})={appear|出現} or cause({appear|出現})={CauseToAppear|顯現}. In the future, important relations regarding entailment and precondition between two concepts will be further explored.

5. Discussion and Conclusion

Levin (2010) pointed out that different studies support positing verb classes of varying grain-sizes, including (a) coarse-grained classification discriminating *manner verb*, *result verb*; (b) medium-grained classification discriminating *motion verbs*, *speaking verbs*, *etc*., with Fillmore’s verb classification being regarded as a representative of medium-grained classification; and (c) fine-grained classification discriminating *run*, which lexicalizes a manner of motion that causes directed displacement towards a goal. Nevertheless, while these classifications are different in grain-size, they are not contradictory for the classification criteria.

In E-HowNet, we carry this viewpoint through the whole construction by first classifying events into causative processes and their corresponding resultant states, *i.e.*, the top two levels of events we mainly discussed in this paper. We then further subdivided these into more than 1200 generic events (*i.e.*, primitives) into a semantic hierarchy framework as a medium-grained event classification. Finally, the near synonyms were attached to each primitive and discriminated by fine-grained features that were integrated in the lexical event structure representation of E-HowNet (abbreviated as LESRE; see Chen *et al.*, 2013). The content and formation of LESRE is shown in Figure 3.

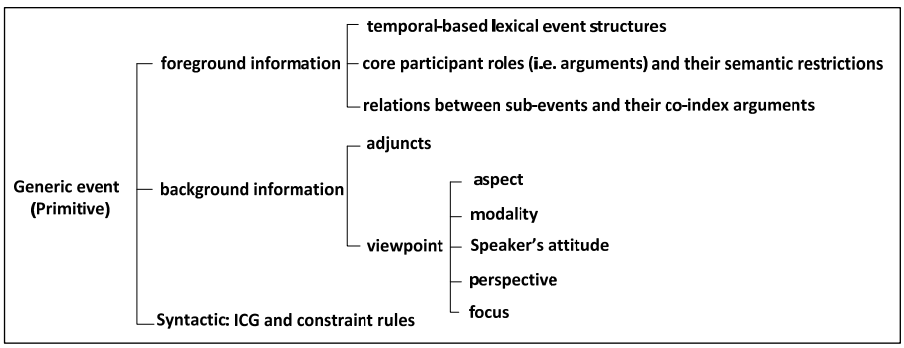


Figure 3. The Content and Formation of LESRE

We believe the varying grain-size classifications provide different semantic and syntactic realization options, such as the coarse-grained classification in which process verbs show considerably more and different argument options than state verbs; further, the idiosyncrasy of each grain-size classification, as well as their interaction, will provide us with advanced knowledge in lexical representation. We will, therefore, continue to complete the LESRE theory in the near future, with the ultimate objective being to establish a completed event classification system that can be applied to both theoretical and computational linguistics. The sketch of different grain-sized event classifications in the E-HowNet construction is shown in Figure 4.

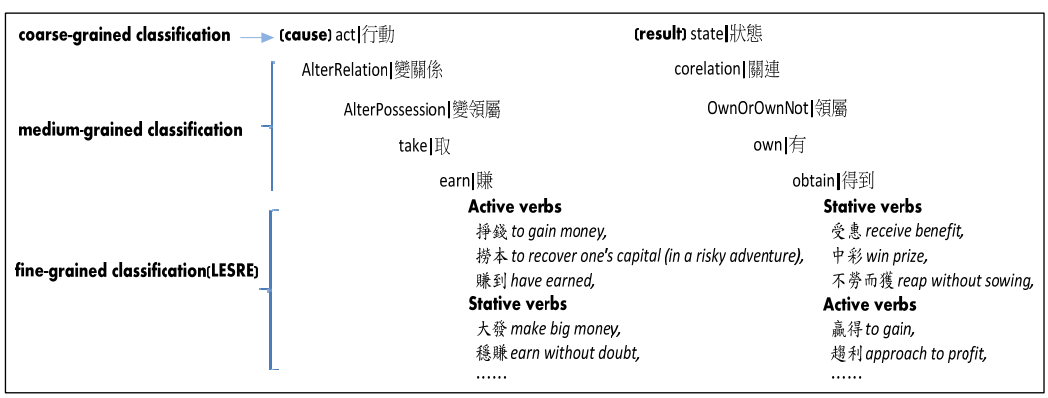


Figure 4. Three Grain-sizes of Event Classification in E-HowNet Construction

Event classification is one of the crucial tasks in lexical semantic representation. Traditionally, researchers have regarded process and state as the two top level events and defined them by counter temporal features and syntactic rules. In this paper, we added cause-result relativity as an auxiliary criterion to discriminate between process and state, and structured about 40,000 Chinese verbs to the two correspondent event classes. All verbs were classified according to their semantic similarity with the conceptual types of the ontology. The process or state type mismatched verbs would have their types adjusted by their PoS or semantic expressions. Furthermore cause-result relations would be linked between corresponding processes and states to bridge the gaps of the dichotomy approach.

We not only aimed to claim the deficiency of dichotomy approach, but also to point out that any discrete event classification system is insufficient to make a clear-cut classification for all verbs, such as synonyms with slightly different semantic focuses. Although misclassification maybe unavoidable, under our framework of event classification, we proposed the remedy of using fine-grained feature expressions to recover erroneous information inherited from the mismatched classification and differentiated the fine-grained semantic differences for near synonyms. The E-HowNet feature expression system is an incremental system, *i.e*., fine-grain features can be added gradually without side effects. Currently, we have resolved the medium-grained classification among 1200 generic event types for about 40,000 Chinese verbs. In the future, we will improve their fine-grained feature expressions to achieve better lexical semantic and syntactic representations.

Acknowledgments

This research was supported in part by National Science Council under Grant NSC 99-2221-E-001-014-MY3.

Reference

Aristotle, (1984), Metaphysics. In Jonathan Barnes (eds.), The Complete Works of Aristotle: The Revised Oxford Translation, Volume 2. Princeton, NJ: Princeton University Press.

Chen, K.-J., *et al*., (2003), *E-HowNet*, CKIP Group, Academia Sinica, http://ehownet.iis.sinica.edu.tw /ehownet.php.

Chen, K.-J., Huang, S.-L., Shih, Y.-Y., & Chen, Y.-J. (2005). Extended-HowNet- A Representational Framework for Concepts. *OntoLex 2005 - Ontologies and Lexical Resources IJCNLP-05 Workshop*, Jeju Island, South Korea.

Chen, K.-J., Huang, S.-L., & Lin, S. (2013), *The Lexical Event Structure Representation of E-HowNet*, technical report, CKIP Group, Academia Sinica. (In preparation).

CKIP group. (1993). *Technical report no.93-05*, Academia Sinica.

Dong, Z., & Dong, Q. (2006). *HowNet and the Computation of Meaning*. World Scientific Publishing Co. Pte. Ltd.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press. http://www-personal.umich.edu/~jlawler/levin.verbs.

Levin, B., (2010). What is the best grain-size for defining verb classes? *Conference on Word Classes: Nature, Typology, Computational Representations, Second TRIPLE International Conference*, Università Roma Tre, Rome, pp. 24-26.

Levin, B., (2011). Verb Classes Within and Across Languages. *Valency Classes Conference, Max Planck Institute for Evolutionary Anthropology*, Leipzig, pp. 14-17.

Meulen, A. G.B. ter, (1983). *The Representation of Time in Natural Language*. In A.G.B. ter Meulen (ed.), Studies in Modaltheoretic Semantics. Dordrecht: Foris.

Meulen, A. G.B. ter, (1995). *Representing Time in Natural Language: The Dynamic Interpretation of Tense and Aspect*. Cambridge, MA: MIT Press.

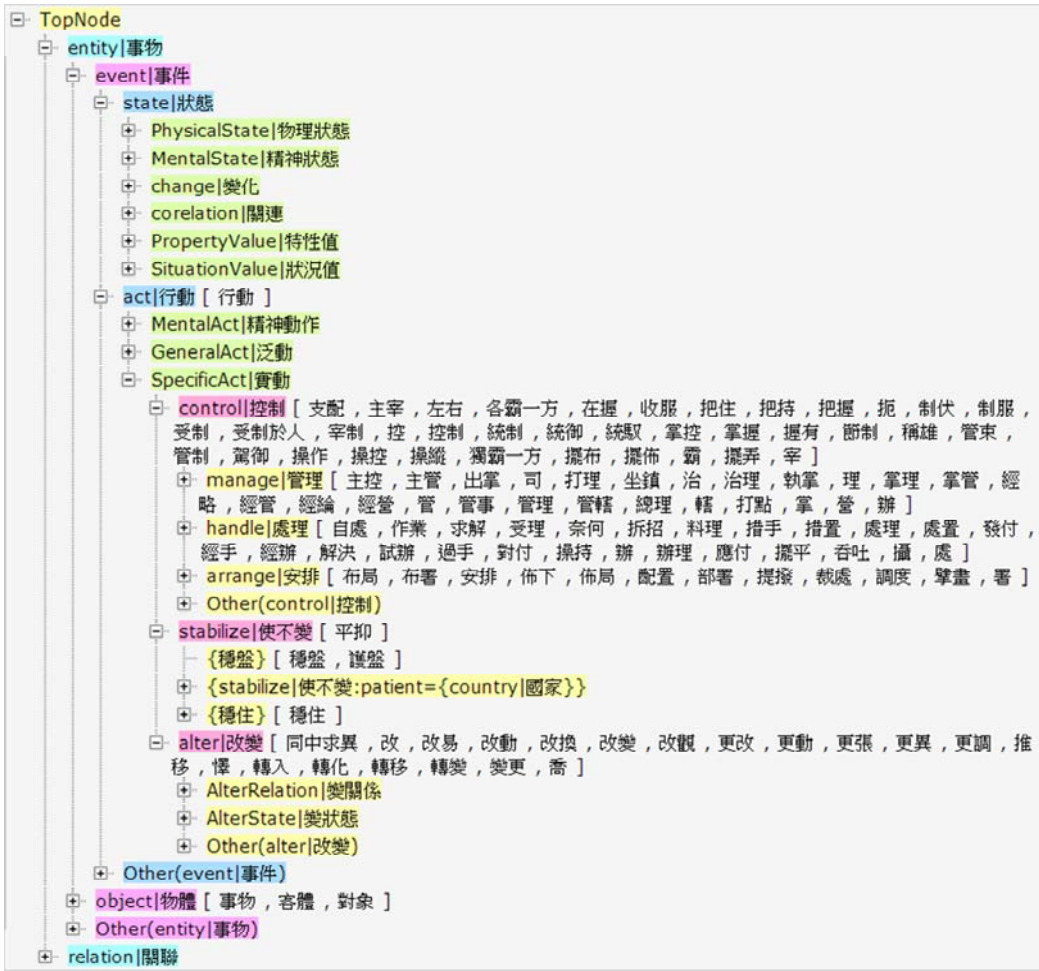
Rappaport, H. M., & Levin, B. (1998). *Building Verb Meanings*. in M. Butt and W. Geuder, eds., The Projection of Arguments, CSLI Publications, Stanford, CA, pp. 97-134.

Rappaport, H. M., & Levin, B. (2005). Change of State Verbs: Implications for Theories of Argument Projection, in N. Erteschik-Shir and T. Rapoport, eds., *The Syntax of Aspect*. Oxford: Oxford University Press, pp. 274-286.

Rosen, S. T., (2003). The Syntactic Representation of Linguistic Events. In L. Cheng & R. Sybesma (eds.), *The 2nd State of the Article Book*. Mouton de Gruyter, Berlin, pp. 323-365. Reprinted from Glot International, 4, 3-10.

Vendler, Z., (1967). *Linguistics in Philosophy*. Ithaca, New York: Cornell University Press, pp. 97-121.

Appendix A: A Snapshot of E-HowNet Ontology



Appendix B: Dual Process-State Type Primitives in the E-HowNet Ontology

	Process	State	→	Dual Process-State	Example
1.	reproduce|生殖	ComeToWorld|問世		ComeToWorld|問世	降生 jiang-sheng ‘born’
2.	CauseToLive|使活	animated|有生命		living|生活	維生 wei-sheng ‘earn a living’
3.	MakeLiving|謀生	alive|活著			
4.	KeepOn|使繼續	GoOn|繼續		KeepOn|使繼續	待續 dai-xu ‘continued’
5.	keep|保持	withstand|抗住		keep|保持	顧全 gu-quan ‘keep in mind’
6.	resume|恢復	BeRecovered|復原		resume|恢復	復甦 fu-shu ‘resuscitate’
7.	stay|停留	situated|處於		stay|停留	獨處 du-chu ‘solitary’
8.	AimAt|定向	facing|朝向		AimAt|定向	迎向 yi-xiang ‘face to’
9.	attract|吸引	attractive|誘人		attract|吸引	迷住 mi-zhu ‘preoccupy’
10.	economize|節省	thrifty|儉		economize|節省	節儉 jie-jian ‘scrimp’
11.	lavish|浪費	extravagant|奢		lavish|浪費	虛靡 xu-mi ‘waste’
12.	ExpressAnger|示怒	angry|生氣		ExpressAnger|示怒	生氣 sheng-qi ‘angry’
13.	forgive|原諒	lenient|寬大		forgive|原諒	寬容 kuan-rong ‘broadminded’
14.	slack|偷懶	lazy|懶		slack|偷懶	混 hun ‘drift along’
15.	pity|憐憫	benevolent|仁		pity|憐憫	心軟 xin-ruan ‘softhearted’
16.	betray|背叛	treacherous|逆		betray|背叛	變節 bian-jie ‘defect’
17.	recreation|娛樂	enjoy|享受		enjoy|享受	自娛 zi-yu ‘amuse oneself’
18.	SeekPleasure|尋歡	enjoy|享受			
19.	None	ill|病態		ill|病態	生病 sheng-bing ‘sick’
20.	None	err|出錯		err|出錯	失誤 shi-wu ‘mistake’
21.	None	lack|缺少		lack|缺少	缺欠 que-fa ‘lack’
22.	None	ServeAsFoil|陪襯		ServeAsFoil|陪襯	相襯 xiang-chen ‘match’

Appendix C: The PoS Features and Semantic Expressions of Examples in the Paper

Examples	PoS	Semantic Expression
購妥 gou-tuo ‘to complete procurement’	active	{buy|買:aspect={Vachieve|達成}}
想起 xiang-qi ‘call to mind’	active	{remember|記得}
記取 ji-qu ‘keep in mind’	active	{remember|記得}
背起來 bei-qi-lai ‘memorize’	active	{remember|記得:aspect={Vachieve|達成}}
念念不忘 nian-nian-bu-wang ‘memorable’	stative	{remember|記得:manner={continuous|連續}}
刻骨銘心 ke-gu-ming-xin ‘be remembered with deep gratitude’	stative	{remember|記得:degree={extreme|極}}
打破 da-po ‘break’	active	{beat|打:patient={x},result={split|破開:patient={x}}}
破裂 po-lie ‘broken’	stative	{FormChange|形變:StateFin={incomplete|缺}}
取悅 qu-yue ‘please’	active	{please|取悅}
喜悅 xi-yue ‘joyful’	stative	{joyful|喜悅}
生氣 sheng-qi ‘angry’	stative	{ExpressAnger|示怒}
發脾氣 fa-pi-qi ‘get angry’	active	{ExpressAnger|示怒}
遇害 yu-hai ‘be murdered’	stative	{kill|殺害}
磨光 mo-guang ‘burnish’	active	{brighten|使亮:means={rub|摩擦}}
擦亮 ca-liang ‘polish’	active	{brighten|使亮:means={wipe|擦拭}}
水亮 shui-liang ‘bright as water’	stative	{bright|明}
光燦 guang-can ‘shining’	stative	{bright|明}
弔死 diao-si ‘hang by the neck’	active	{kill|殺害:means={coil|纏繞}}
往生 wang-sheng ‘pass away’	stative	{die|死}
行刑 xing-xing ‘execute’	active	{punish|處罰:means={kill|殺害}}
處決 chu-jue ‘put to death’	active	{punish|處罰:means={kill|殺害}}
受刑 shou-xing ‘be put to torture’	stative	{punish|處罰:domain={police|警}}
伏法 fu-fa ‘be executed’	stative	{punish|處罰:means={kill|殺害}}

板起(臉) ban-qu-(lian) ‘put on a stern expression’	active	{austere|冷峻}
正色 zheng-se ‘with a stern countenance’	active	{austere|冷峻}
凝重 ning-zhong ‘serious’	stative	{austere|冷峻}
不苟言笑 bu-gou-yan-xiao ‘serious in speech and manner’	stative	{austere|冷峻}
求生 qiu-sheng ‘seek to survive’	active	{living|生活}
度日 du-ri ‘subsist’	active	{living|生活}
生存 sheng-cun ‘exist’	stative	{living|生活}
在世 zai-shi ‘be living’	stative	{living|生活}
一息尚存 yi-xi-shang-cun ‘be still alive’	stative	{living|生活}
累病 lei-bing ‘sick from overwork’	stative	{ill|病態:cause={tired|疲乏}}
驚退 jing-tui ‘frighten off’	active	{frighten|嚇唬: patient={x}, result={leave|離開:agent={x}}}
看中 kan-zhong ‘take fancy to’	active	{FondOf|喜歡}
喜愛 xi-ai ‘love’	active	{FondOf|喜歡}
酷愛 ku-ai ‘ardently love’	active	{FondOf|喜歡:degree={extreme|極}}
熱衷 re-zhong ‘be addicted to’	active	{FondOf|喜歡}
癡情 chi-qing ‘be infatuated’	stative	{FondOf|喜歡:manner={mad|瘋痴}}
興致盎然 xing-zhi-ang-ran ‘full of interest’	stative	{FondOf|喜歡:cause={interesting|趣}}