Causal Web Determination from Texts

Chaveevan Pechsiri¹, Narongdech Keeratipranon¹, and Intaka Piriyakul²

¹ College of Innovative Technology and Engineering, Dhurakij Pundit University, Bangkok, Thailand
² Faculty of Social Sciences, Srinakharinwirot University, Bangkok, Thailand

Email: {chaveevan.pec, narongdech.ken}@dpu.ac.th, intaka@hotmail.com

Abstract—The research aim is to determine a causal web from downloaded guru web-board documents. The causal web which benefits a diagnosis service assistant of a problem-solving system consists of several cause-effect pair sequences where each cause-effect pair has a cause-effect relation and the last cause-effect pair of each cause-effect pair sequence has the same effect concept. Each causative/effect concept is expressed by an elementary discourse unit or a simple sentence. The research has three problems; how to determine the cause-effect pair with an overlap problem between a causative-verb concept set and an effect-verb concept set, how to determine cause-effect pair sequences including causative/effect boundary determination, and how to determine the causal web on the extracted cause-effect pair sequences without redundant sequences. We use a word co-occurrence to represent a sentence’s event/state with a causative/effect concept. We then propose using a self-Cartesian product on a collected causative/effect concept is expressed by an elementary discourse unit, template matching technique to determine the causal web without the redundancy. The research result has a high percentage correctness of the causal web determination.

Index Terms—word co-occurrence, elementary discourse unit, template matching

I. INTRODUCTION

Determination of a causal web from downloaded documents on the guru web-boards is a challenge where a causal web is a metaphor that emphasizes the interconnectedness of casual components in a population including direct causes and indirect causes comprising causal webs [1]. The causal web expression of a certain domain benefits a diagnosis service assistant of a problem-solving system in that domain. The aim of this research is then to determine the causal web from the downloaded guru documents, particularly on the car-problem guru web-boards (i.e. AskGuru [2], MThai [3], etc.).

Regard to the causal web explanation by [1], the causal web of our research consists of several cause-effect pair sequences expressed on the documents where each cause-effect pair has a cause-effect relation between one or more event/state expressions with causative concepts and one or more event/state expressions with effect concepts.

The event/state expression with the causative/effect concept is based on an Elementary Discourse Unit (EDU) defined as a simple sentence / a clause by [4]. Moreover, the last cause-effect pairs from all cause-effect pair sequences have events/states with the same effect concept to construct the causal web as shown in Fig. 1 where a node represents an event/state expressed by EDU with a causative/effect concept and a link represents a cause-effect relation of a cause-effect pair from a cause node to an effect node pointed by an arrow. Fig. 1 also shows the alternative causes for assisting automatic human reasoning in diagnosis of the car problems/ symptoms.

Each EDU (see Fig. 2) is expressed by the general Thai linguistic expression after stemming words and eliminating stop words where NP1 and NP2 are noun phrases, VP is a verb phrase, V is a non-terminal verb expression, Verb_ment is a strong-verb concept set, Verb_weak is a weak-verb concept set needed more information, Adv is an adverb concept set, Noun is a noun concept set, and Adj is an adjective concept set. All concepts of these concept sets are based on WordNet [5] and Thai Encyclopedia [6] after translating from Thai to English by Lexitron [7]. For example: Example1 (see Fig. 3) shows the cause-effect relation occurs between an EDU1 expression with an effect concept and both EDU2 and EDU3 expressions with causative concepts as shown in the following cause-effect pair expression.

EDU2 ∧ EDU3: Cause → EDU1: Effect

Example2 (see Fig. 4 where a [..] symbol means ellipsis) contains the cause-effect pair sequence as shown in the following where EDU2, EDU3 and EDU4 express the causative concepts and also the effect concepts.

EDU1: Cause → EDU2: Effect
EDU2: Cause → EDU3: Effect
EDU3: Cause → EDU4: Effect
EDU4: Cause → EDU5: Effect

Figure 1. Causal web of car problems, i.e. “A car won’t start.”
Figure 2. Linguistic pattern after stemming words and stop word removal

Figure 3. Show cause-effect relation between 2 causative-concept EDUs and 1 effect-concept EDU

Figure 4. Show a sequence of cause-effect EDU pairs

There are several techniques, [8]-[14], having been applied for determining the cause-effect/causality/causation relation from texts (see Section II). However, the Thai documents have several specific characteristics, such as zero anaphora or the implicit noun phrase, without word and sentence delimiters, and etc. All of these characteristics are involved in three main problems (see Section III). The first problem is how to determine a cause-effect pair from an EDU pair having a cause-effect relation with the overlap problem between-a causative-verb concept set (V\text{cause}) and an effect-verb concept set (V\text{effect}). Where V\text{strong} = V\text{cause} \cup V\text{effect}. The second problem is how to determine the cause-effect pair sequences including the EDU boundary determination with either the causative-event/state concepts or the effect-event/state concepts. And the third problem is how to determine the causal web from the extracted cause-effect pair sequences with the problem of redundant cause-effect pair sequences. Regarding all of these problems, we develop a framework which combines machine learning and the linguistic phenomena to learn the several EDUs having the cause-effect relation on the downloaded documents. Therefore, we collect a word co-occurrence (called wrdCo) pattern from each EDU expression after stemming words and eliminating stop words into a wrdCo set (called WC). Each wrdCo expression/element has the pattern as shown in the following to represent an EDU occurrence with a causative-event/state concept or an effect-event/state concept where each causative-effect concept is expressed by VP of an agent expressed by NP1.

wrdCo Pattern: \(w_1 + w_2 + w_3 + w_4\)

where \(w_1\) is a head noun of NP1 and has a null value if NP1 is ellipsis; \(w_2, w_3,\) and \(w_4\) exist on VP having \(w_2 \in (V\text{strong}\setminus V\text{weak})\); \(w_2\) and \(w_4\) are a word sequence right after \(w_2\) and have a null value if they don’t exist. And \(w_3,w_4\) are Word.

We then propose using a self-Cartesian product on the collected WC with concepts along with the Naive Bayes (NB) learning technique [15] to determine WCP which is a wrdCo ordered pair set having the cause-effect relation as follow: WCP\text{ce} = \{wcpair1, wcpair2, ..., wcpair\text{last}\}; and each WCP\text{ce} element (wcpair\text{last}) has the cause-effect relation between two wrdCo expressions, one wrdCo expression with a causative-event/state concept and the other one with an effect-event/state concept. We apply an experimental Event Rate (ER) [16] between two related event/state-concept occurrences for the verb categorization into verb groups/sets (see Section III.A). WCP\text{ce} and categorized verb groups are used to identify each EDU pair with the cause-effect relation without the influence of the overlap between \(V\text{cause}\) and \(V\text{effect}\) (\(V\text{cause} \cap V\text{effect} \neq \emptyset\)). We then extract the wrdCo pair sequence as the cause-effect pair sequence from the documents by using WCP\text{ce} including the categorized verb groups to solve the wrdCo/EDU boundary with the causative concept or the effect concept (see Section III.B). And we apply a dynamic template matching the last cause-effect pair of each extracted cause-effect pair sequence to discard the redundancy problem (see Section III.C).

Our research is organized into 5 sections. In Section II, Related Works are summarized. Problems in determining the causal web from texts are described in section III and Section IV shows a framework of Causal Web Determination from Texts. In Section V, we evaluate and conclude our model.

II. RELATED WORKS

Several strategies, [8]-[14], have been proposed to determine the cause-effect relation from texts without considering the cause-effect pair sequence except [13]. Reference [11] applied Integer Linear Programming to learn the causal relation on a sentence from annotated
verb-noun pairs based on FrameNet, WordNet and linguistic features. Reference [8] identified the causal relation between two adjacent sentences by using Support Vector Machine to learn several features from the two sentences as causal volition, the verb class, verbal semantic attributes, the connective marker, and modality. Reference [10] determined a predicate pair, which is an event pair from two sentences, having a causality by measuring a cause-effect association based the point wise mutual information with minimally supervised approach. Reference [9] applied verb-pair rules resulted by machine learning techniques to extract the causality from several EDUs including the cause/effect EDU boundary determination without the cause-effect pair sequence consideration. Reference [12] proposed the Restricted Hidden Naïve Bayes model to the lexico syntactic pattern on a sentence to learn and extract the causality with the contextual, syntactic, positional, and connective features from the English documents. Reference [14] extracted the causal relation within one or two sentences by using the linguistic rules based along with Bayesian inference to reduce the number of pairs produced by ambiguous patterns. Reference [13] applied the Granger causality model with features, i.e. N-words, topics, sentiments and etc., to detect cause - effect relationships from text for a time series and also used a neural reasoning algorithm to construct chain of cause and effect pairs as an explanation result with 57% accuracy.

However, most of the previous works on the cause-effect relation are based on event/state expressed by either NP or VP features mostly existing in one/two sentences without considering the cause-effect pair sequences enclosed in the causal web. However, there are few works on determining cause-effect pair sequences as causal pathways.

III. PROBLEMS OF DETERMINING CAUSAL WEB

A. How to Determine a Cause-Effect Pair Having V_cause and V_effect Overlap

To determine the cause-effect pair having the cause-effect relation among several EDUs, we apply the wrdCo expression as mention in section I to represent an EDU’s event/state concept along with NP1 as an EDU’s agent. We use V_cause and V_effect collected from an annotated learning corpus to identify the EDU occurrence with the causative-event/state concept or the effect-event/state concept if V_cause \cap V_effect = \emptyset. However, there is the V_cause and V_effect overlap problem in our documents as shown in Example2 having V_cause \cap V_effect \neq \emptyset on EDU2, EDU3 and EDU4. Therefore, the categorized verb groups and WCP_ce including the String Matching technique are integrated to identify a wrdCo pair of an EDU pair as the cause-effect pair having the cause-effect relation from the documents. Where WCP_ce is determined by the NB-learning probabilities of wrdCo concept pairs with the CauseEffectRelation class from the annotated learning corpus to the self-Cartesian product of the collected wrdCo set or WC including concept expressions after stemming words and eliminating stop words from the testing corpus. We apply ER to measure the frequencies of the \(v_c\) occurrences (\(v_c \in \text{V}_{\text{strong}}\)) and the (\(v_c + w_d\) or \(v_c \in \text{V}_{\text{weak}}\) and \(w_d \in \text{Word}\) on the linguistic pattern) as causative-event/state concepts and/or effect-event/state concepts on the annotated learning corpus for the verb categorization into three verb groups/sets, a cause group (VC), an cause/effect group (UCE), and an effect group (VE) as follow.

\[
\text{ER-of-}v_c = \frac{\text{theNumberOf } v_c}{(\text{theNumberOf } v_c + \text{theNumberOf } v_{\text{weak}})}
\]

(1)

\[
\text{ER-of-}v_c = \frac{\text{theNumberOf } v_{\text{weak}}}{\text{theNumberOf } v_c + \text{theNumberOf } v_{\text{weak}}}
\]

(2)

where \(v_c\) is \(v_c\) with a causative-event/state concept; \(v_{\text{weak}}\) is \(v_c\) with an effect-event/state concept;

\[
\text{ER-of-}v_{\text{weak}},_{\text{wd}} = \frac{\text{theNumberOf } v_{\text{weak}},_{\text{wd}}(\text{theNumberOf } v_{\text{weak}},_{\text{wd}} + \text{theNumberOf } v_{\text{weak}},_{\text{wd}})}{\text{theNumberOf } v_{\text{weak}},_{\text{wd}}}
\]

(3)

\[
\text{ER-of-}v_{\text{weak}},_{\text{wd}} = \frac{\text{theNumberOf } v_{\text{weak}},_{\text{wd}}}{\text{theNumberOf } v_{\text{weak}},_{\text{wd}}} + \text{theNumberOf } v_{\text{effect}},_{\text{wd}}
\]

(4)

where \(v_{\text{weak}},_{\text{wd}}\) is \(V_{\text{weak}},_{\text{wd}}\) with a causative-event/state concept; \(v_{\text{weak}},_{\text{wd}}\) is \(V_{\text{weak}},_{\text{wd}}\) with an effect-event/state concept.

Equation (1)-(4), the verb expressions can be categorized by their ER values into VC (if ER-of-\(v_c\) or ER-of-\(v_{\text{weak}},_{\text{wd}}\) \(\geq 0.9\), VE (if ER-of-\(v_{\text{weak}}\) or ER-of-\(v_{\text{weak}},_{\text{wd}}\) \(\geq 0.9\), or otherwise VCE.

B. How to Determine Cause-Effect Pair Sequences Including Boundary Determination

The cause-effect pair sequence sometimes contains an EDU boundary with the causative concept as shown in Fig. 5 as Example3 having EDU1 and EDU2 as the causes of EDU3, and also an EDU boundary with the effect concept as shown in Fig. 5 as Example4 having EDU4, EDU5, and EDU6 as the effects of EDU3.

Example3:

EDU1: “[we] smell” (“We sometimes park a car.”) ([((we/>wNP1) (we/parked)/V (to become)/vapor)/VP)"

EDU2: “... /boil /to become /vapor” (“and we carelessly leave a light on overnight.”) ([we/>wNP1] (we/carelessly /left)/V (light on)/V (overnight)/V)

EDU3: “... /rub /the battery’s power drains.” ([rubbattery powerNP1] (rub /the battery’s powerNP1)/VP)"

EDU4: “... /smell /light on /for a long time” ([smell/>wNP] (smell /light on)/V (for a long time)/V)

Example4:

EDU1: “[we] drive” (“We pedal the brakes all the time down the mountain.”) ([we/>weNP1] (we/pedaled)/V (to become)/vapor /all time /down /mountain)/VP)

EDU2: “... /drive /for a long time” (“we /drive /the brakes /for a long time”) ([drive/>wNP1] (drive /the brakes /for a long time)/VP)

EDU3: “... /ollie /to sky” (“The brakes then have high heat.”) ([ollie/>wNP1] (ollie /to sky)/V (high heat)/V)

EDU4: “... /burn /... /burn /the brakes /to sky” (“Cause the brakes pads to be burnt.”) ([burn/>wNP1] (burn /the brakes padNP1)/VP)

EDU5: “... /become /... /smell” (“and the brake pads have a burning smell.”) ([become/>wNP1] (become /the brake padsNP1)/VP)

EDU6: “... /burn /... /boil” (“The brake oil sometimes boils into vapor.”) ([burn/>wNP1] (burn /the brake oilNP1)/VP)

Where a [...] symbol means ellipsis.

Figure 5. Show a causative-EDU boundary on Example3 and an effect-EDU boundary on Example 4.
Example 3 on Fig. 5 shows a cause-effect pair sequence having EDU1 and EDU2 as a causative-EDU boundary as follows:

\((\text{EDU1} \land \text{EDU2}) \rightarrow \text{EDU3} \land \text{EDU4} \rightarrow \text{EDU5} \)]

Example 4 on Fig. 5 also shows a cause-effect pair sequence having EDU4, EDU5, and EDU6 as an effect-EDU boundary as follows:

\((\text{EDU1}) \rightarrow \text{EDU2} \land \text{EDU3} \rightarrow \text{EDU4} \land \text{EDU5} \land \text{EDU6}) \rightarrow \text{EDU7}) \)

Therefore, we solve the event/state boundary with either causative concept or effect concept by using the categorized verb groups by ER along with WCP_ε through String Matching.

C. How to Determine Causal Web with Redundancy Problem

There is variety of the extracted cause-effect pair sequences with some redundant sequences from the documents. Therefore, we apply a dynamic template based on concept expressions for String Matching in the last cause-effect pair of each extracted cause-effect pair sequence to eliminate the redundant cause-effect pair sequences which match patterns hold by the template. Whereas, a simple template (which is a concept pattern of the last cause-effect pairs of the extracted cause-effect pair sequences) consists of a wrdCo ordered pair with concept expression as follow:

SimpleTemplate

\((\text{wrdCo}_\text{cause}: \text{ConceptExpression}) \land \text{wrdCo}_\text{effect}: \text{ConceptExpression})\)

where \(\text{wrdCo}_\text{cause}\) is a word co-occurrence with a causative- event/state concept; \(\text{wrdCo}_\text{effect}\) is a word co-occurrence with an effect-event/state concept; \(\text{wrdCo}_\text{cause},\text{w}_2 / \text{wrdCo}_\text{cause},\text{w}_3 \in \text{VC/VE}; \text{and wrdCo}_\text{cause},\text{w}_2 \rightarrow \text{wrdCo}_\text{cause},\text{w}_3 \in \text{VC/VCE}.\)

The simple template will discard only the cause-effect pair sequence having the last cause-effect pair being the same as the template pattern. With regard to several extracted cause-effect pair sequences, the simple template pattern can be automatically adjusted to contain several wrdCo ordered pairs having only the same \(\text{wrdCo}_\text{effect}\): ConceptExpression with different \(\text{wrdCo}_\text{cause}\): ConceptExpression as follow:

DynamicTemplate

\((\text{wrdCo}_\text{cause},i: \text{ConceptExpression}) \land \text{wrdCo}_\text{effect}: \text{ConceptExpression})\)

where \(i=1, 2, n\); \(n\) is an integer; \(\text{wrdCo}_\text{effect},\text{w}_2 \land \text{wrdCo}_\text{effect},\text{w}_3 \in \text{VE}; \text{and wrdCo}_\text{cause},\text{w}_1 \rightarrow \text{wrdCo}_\text{cause},\text{w}_3 \in \text{VC/VCE}.\)

IV. A FRAMEWORK OF DETERMINING CAUSAL WEB

The causal web determination consists of six steps; Corpus Preparation, Categorized Verb Group Determination, Learning wrdCo Pair with CauseEffectRelation, Extraction of wrdCo ordered pairs having CauseEffectRelation, Extraction of Cause-Effect Pair Sequences, and Causal Web Determination (see Fig. 6).

A. Corpus Preparation

This step is to prepare an EDU corpus of car-problem documents downloaded from the car guru web-boards. The step involves using Thai word segmentation tools [17], Named Entity recognition [18], and EUD Segmentation [19] to provide the 2500 EDUs’ corpus. The corpus included stemming words and the stop word removal is separated into 3 parts; the first 1000-EDUs’ part as the annotated learning corpus for categorizing verb groups by ER and learning the wrdCo pairs with CauseEffectRelation or NonCauseEffectRelation. The second 1000-EDUs’ part is for testing as extracting and collecting of the WCP_ε set. The third 500-EDUs’ part is for testing as extracting the cause-effect pair sequences used for determining the causal web. We then semi-automatically annotate the wrdCo pattern with concepts along with the categorized verb groups as VC, VCE, and
VE on each EDU of the learning corpus (see Fig. 7). All wrdCo concepts are referred to WordNet [5] after the Thai-English translation by Lexitor [7].

### B. Categorized Verb Group Determination

Regard to the annotated learning corpus, we collect each VP tag containing a group property as Group=VC, Group=VCE, or Group=VE for the ER determination. Equation (1)-(4) on Section III.A, the verb expression can be categorized by ER value into three verb groups/sets, VC, VCE, and VE (see Fig. 8), with regard to the following ER-rule where \( w_d \) is a word occurrence annotated by the \( w \) tag (see Fig. 7).

**ER-rule:**

- If ER-of-\(v_{sd} \geq 0.9 \) or ER-of-\( (v_{sd},w_d) \geq 0.9 \) then \( v_{sd} \in VC \) or \( v_{sd}(w_d) \in VC \) respectively
- Else If ER-of-\(v_{sd} \geq 0.9 \) or ER-of-\( (v_{sd},w_d) \geq 0.9 \) then \( v_{sd} \in VCE \) or \( v_{sd}(w_d) \in VCE \) respectively
- Else \( v_{sd} \in VE \) or \( v_{sd}(w_d) \in VE \) respectively

**Figure 8. Categorized verb group: VC, VCE, and VE**

### C. Learning wrdCo Pairs with CauseEffectRelation

This step is the NB learning [15] the feature set of wrdCo-concept pairs with the CauseEffectRelation/ nonCauseEffectRelation class on several two adjacent EDUs with CEREId/nonCEREId annotation from the corpus preparation step (Section IV.A) after stemming words and eliminating stop words. The learning results of this step by using Weka [20] are the probabilities of the wrdCo pairs with concept expressions from the annotated learning corpus as shown in Table I.

<table>
<thead>
<tr>
<th>wrdCoPair with Concept Expression:</th>
<th>CauseEffectRelation</th>
<th>NonCauseEffectRelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Park(4,car,LeaveBehind)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FailToStart(car)</td>
<td>0.02367942</td>
<td>0.05732484</td>
</tr>
<tr>
<td>Leave(4,lightOn)</td>
<td>0.02367942</td>
<td>0.00636943</td>
</tr>
<tr>
<td>Loose(battTerminal)</td>
<td>0.00884956</td>
<td>0.00429185</td>
</tr>
<tr>
<td>NotStartCar(electricSystem)</td>
<td>0.008910747</td>
<td>0.00636943</td>
</tr>
</tbody>
</table>

### D. Extraction of wrdCo Ordered Pairs Having CauseEffectRelation

We collect all wrdCo expressions from the 1000EDU testing corpus including concepts from the corpus preparation into WC\(_{\text{concept}}\) (which is WC including concept expressions) used for the self-Cartesian product as WC\(_1 \times WC_2\) where WC\(_1\) is WC\(_{\text{concept}}\) used as the causative-event/state concept set; WC\(_2\) is WC\(_{\text{concept}}\) used as the effect-event/state concept set. The result of the self-Cartesian product on WC\(_{\text{concept}}\) is WCP\(_{ce}\) which is the wrdCo order pair set (including the concept expression of each wrdCo order pair) having the cause-effect relation through NB as shown in (5) along with the probabilities of the wrdCo pair features with CauseEffectRelation and nonCauseEffectRelation from the learning corpus of section IV.C.

\[
\begin{align*}
\text{weOrdpair } \_ \text{Class} &= \arg \max \text{ P} (\text{class} | \text{wcpair}_k) \\
\text{classClass} &= \arg \max \text{ P}(\text{wcpair}_k | \text{class})(\text{P}(\text{class}) \\
\text{wcpair} &\in \text{WC}_1 \times \text{WC}_2; k = 1,2,\ldots; \text{num} \\
\text{num} &\text{ is the number of WC}_j \times \text{WC}_k \\
\text{Class} &= \{\text{CauseEffectRelation, nonCauseEffectRelation}\}
\end{align*}
\]

**Figure 9. Cause-Effect pair sequence extraction algorithm**
F. Causal Web Determination

In regard to the downloaded documents, the research concerns only two different template patterns as shown in the following.

**DynamicTemplate1:**

\[(\text{wrdCo}_{\text{cause}},:\text{ConceptExpression}) \rightarrow \text{FailToStart(\text{car})}\]

**DynamicTemplate2:**

\[(\text{wrdCo}_{\text{cause}},:\text{ConceptExpression}) \rightarrow \text{drop(\text{enginePower})}\]

We search for matching between the last wrdCo ordered pair as the last cause-effect pair of each extracted cause-effect pair sequence (Sequence\(i\), \(a=1,2,...,\text{lastSequence}\)) and the specified templates, DynamicTemplate1 and DynamicTemplate2, where each template consists of one wrdCo\(_{\text{cause}-i}\) and one wrdCo\(_{\text{effect}}\). If the exact match of either DynamicTemplate1 or DynamicTemplate2 is found, the extracted cause-effect pair sequence containing the match of the last cause-effect pair is discarded. If the match occurs only the concept expression of wrdCo\(_{\text{effect}}\) on either DynamicTemplate1 or DynamicTemplate2, this will result in the adjustment of either DynamicTemplate1 or DynamicTemplate2 respectively with wrdCo\(_{\text{cause}(i+1)}\) having ConceptExpression from the Causative\_wrdCoConcept expression of the matched Effect\_wrdCoConcept expression on the last cause-effect pair of the extracted cause-effect pair sequence.

V. EVALUATION AND CONCLUSION

There are three evaluations of the proposed research being evaluated by three expert judgments with max win voting: the first evaluation is the extraction of the WCPE set from 1000 EDUs testing corpus. The second evaluation is the cause-effect pair sequence extraction from the 500 EDUs testing corpus and the third evaluation is the causal web determination from the extracted cause-effect pair sequences. The first and second evaluations are based on the precision and the recall with tenfold cross validation whilst the third evaluation is the percentage of correctness (see Table II).

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Precision</th>
<th>Recall</th>
<th>%Correctness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraction of WCPE(_{\text{cor}}) (from 1000EDU-Testing Corpus)</td>
<td>0.88</td>
<td>0.62</td>
<td>-</td>
</tr>
<tr>
<td>Extraction of Cause-Effect Pair Sequences (from 500EDU-Testing Corpus)</td>
<td>0.90</td>
<td>0.65</td>
<td>-</td>
</tr>
<tr>
<td>Causal Web Determination</td>
<td>-</td>
<td>-</td>
<td>85%</td>
</tr>
</tbody>
</table>

The reasons of having low recalls in both extracting the WCPE\(_{\text{cor}}\) set and the cause-effect pair sequences from the documents are 1) the noun clause occurrence as the direct object of the complex sentence expresses the effect event concept as “that the electric system is not fully functional” in EDU2 of Example6 on Fig. 10. 2) two cause-effect pair sequences express on the documents with sharing the last cause-effect pair as shown in Example6 of Fig 10. Example6 on Fig. 10 consists of two consecutive cause-effect pair sequences as shown in the following with sharing the last cause-effect pair (EDU5) where the first cause-effect pair sequence cannot be extracted.

The first cause-effect pair sequence:  
EDU1:Cause \(\rightarrow\) EDU2:Effect  
EDU3:Cause \(\rightarrow\) EDU4:Effect  
EDU2:Cause \(\rightarrow\) EDU5:Effect  
EDU4:Cause \(\rightarrow\) EDU5:Effect

Example5:

EDU1: “...”  
EDU2: “...”  
EDU3: “...”  
EDU4: “...”  
EDU5: “...”

Moreover, An ellipse of EDU having wrdCo\(_{\text{effect}},2\) or wrdCo\(_{\text{effect}},2\) \(w\) \(s\) \(VCE\) as in EDU2 of Example7 on Fig. 11 results in the redundant cause-effect pair sequences which effect to the % of correctness of the causal web determination.

Example7:

EDU1: “...”  
EDU2: “...”  
EDU3: “...”  
EDU4: “...”  
EDU5: “...”

Figure 10. Show examples of having low recalls

Hence, the research contributes the methodology to determine the causal web for diagnosis service assistant, particularly diagnosis common car symptoms, for car troubleshooting and maintenance to people on social networks. Finally, our research methodology can also be applied to another area, e.g. environmental business, for representing a business activity and/or a human behavior to an environmental change whereas environmental business focuses on how to operate in an eco-friendly way and still remain profitable.

CONFICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Chaveevan Pechsiri is the main author of this work. Chaveevan Pechsiri is responsible for the compilation and drafting of literature, methods, results, evaluation and discussion within this work. Chaveevan Pechsiri also designed the systems overview. Narongdech Keeratipranon is responsible for conducting the machine
learning technique experiments. Intaka Piriyakul is responsible for corpus preparation for this study.

REFERENCES


