The Robotic Surgery Procedural Framebank

Marco Bombieri[1], Marco Rospocher[2], Simone Paolo Ponzetto[3], Paolo Fiorini[1]

[1]Department of Computer Science, University of Verona (Italy),
[2]Department of Foreign Languages and Literatures, University of Verona (Italy)
[3]Data and Web Science Group, University of Mannheim (Germany)
{marco.bombieri_01, marco.rospocher, paolo.fiorini}@univr.it, simone@informatik.uni-mannheim.

Abstract

Robot-Assisted minimally invasive surgery is the gold standard for the surgical treatment of many pathological conditions, and several manuals and academic papers describe how to perform these interventions. These high-quality, often peer-reviewed texts are the main study resource for medical personnel and consequently contain essential procedural domain-specific knowledge. The procedural knowledge therein described could be extracted, e.g., on the basis of semantic parsing models, and used to develop clinical decision support systems or even automation methods for some procedure's steps. However, natural language understanding algorithms such as, for instance, semantic role labelers have lower efficacy and coverage issues when applied to domain others than those they are typically trained on (i.e., newswire text). To overcome this problem, starting from PropBank frames, we propose a new linguistic resource specific to the robotic-surgery domain, named Robotic Surgery Procedural Framebank (RSPF). We extract from robotic-surgical texts verbs and nouns that describe surgical actions and extend PropBank frames by adding any of new lemmas, frames or role sets required to cover missing lemmas, specific frames describing the surgical significance, or new semantic roles used in procedural surgical language. Our resource is publicly available and can be used to annotate corpora in the surgical domain to train and evaluate Semantic Role Labeling (SRL) systems in a challenging fine-grained domain setting.

Keywords: semantic role labeling, robotic-assisted surgical procedures, annotated dataset

1. Introduction

Surgical practice has steadily improved in recent decades thanks to the constant support of the approaches made available by observational science. Examples of these approaches can be found in the rigorous methodologies employed in the research and development of new clinical protocols, drugs or surgical instruments. To further support this constant improvement, it was necessary to introduce the so-called Electronic Health Records (EHR) and hand-written operative notes in the operating theatres, which led to the proliferation of electronic documents concerning operational notes and clinical cases written in natural language (Wang et al., 2013). These reports are very schematic and concise in form and language, and they are targeting expert clinician, highlighting particular cases or anomalies that occurred during the surgery. However, in practice they are often written by medical trainees and missing essential details (Nyamulani and Mulwafu, 2018). Academic papers, online tutorials, and robotic surgery textbooks are also continuously being published and peer-reviewed. Compared to EHR, these resources present the procedure more organically, with a less schematic and more detailed prose-style treatment (Bombieri et al., 2021). These types of texts are very valuable resources because they encode the so-called *surgical procedural knowledge* together with background and outline information. The surgical procedural knowledge is the one possessed by an intelligent agent (in surgery, a surgeon or a surgical robot) able to perform a task (a surgical intervention). Typically, the description of a procedure details a set of surgical actions linked together temporally and causally. Many of these actions are enriched with information providing additional details, such as the tool to use, the surgical technique to follow, the anatomical parts to operate, a set of spatial or temporal attributes and a purpose. The following sentence is an example of procedural knowledge in surgery (here annotated with pseudo-semantic roles in brackets):

(**EX1**) [TEMP: *After dissection of the hilar vessels*], [ANATOMY: *Gerota's fascia*] [ACTION: *is incised*] [PURPOSE: *to expose the tumor and the surrounding renal capsule*].

This in contrast to other sentences that encode knowledge but not of a procedural type, i.e., they do not express an action executable by an intelligent actor:

(**EX2**) *The renal fascia separates the adipose capsule of kidney from the overlying pararenal fat.*

Developing Natural Language Understanding (NLU) methods for mining procedural surgical knowledge is crucial for improving situation awareness modules and knowledge-based decision-making techniques (Demner-Fushman et al., 2009). Mining algorithms can also summarize all the procedural information spread throughout the texts in a structured and organic form usable as a study resource by medical students, or for the automatic generation of operative guidelines, or for drafting safety checklist.

To extract relational information based on actions and actors involved in them (like that of EX1), Semantic Role Labeling (SRL) techniques (Gildea and Jurafsky, 2002) have shown to be a promising and viable solution

(Rebmann and van der Aa, 2021; Sadhu et al., 2021). These methods are based on shallow semantic parsing and produce predicate-argument structures of sentences. In most semantic theories, predicates are verbs, verbal nouns, and some other verb forms. They are based on lexical resources such as PropBank (Palmer et al., 2005) and FrameNet (Baker et al., 1998). PropBank uses both proto-roles and verb-specific semantic roles, while FrameNet uses semantic roles that are specific to a general semantic idea, shared by the (possibly) many lexical units (e.g., verbs) evoking it.

SRL methods based on PropBank are used in numerous NLU applications, such as conversation analysis (Xu et al., 2021), video understanding (Sadhu et al., 2021), information extraction and ontology population (Humphreys et al., 2021), mining of event logs written in natural language (Rebmann and van der Aa, 2021) or automatic image captioning (Chen et al., 2021). However, the performance of the current SRL systems on out-of-domain testing examples is often very poor (Do et al., 2015). This is because PropBank annotation focus on general-purpose, newswire texts and do not fully cover specific domains, such as, in our case, the surgical one. With this work we aim at filling this gap and extend PropBank so that its frames are suitable for representing the semantic roles typically required in the procedural surgical domain. As presented later, semantic frames in this domain are significantly different from the ones of general domain English.

Our key contribution is the public release of a new linguistic resource, which extends PropBank with frames describing actions and participants used in the robotic-assisted surgical domain. This frame repository is essential for adapting SRL methods designed for other domains to the procedural robotic-surgical one. We focus on the context of robotic surgery because it includes all traditional surgical actions plus others that are ad-hoc for it. Furthermore, surgical robotic systems are increasingly pervasive in operating theaters thanks to the superior characteristics (e.g., enhanced stereo view and improved dexterity) that allow the surgeon to perform interventions with a better outcome for patients (Bombieri et al., 2020).

We organize the paper as follows: Section 2. presents a survey of previous works on building semantic framebanks in specific domains or languages. In Section 3., we investigate how the usage of lemmas differ between our domain and the general one. Furthermore, we identify domain keywords not already encoded in state-of-the-art resources, thus adapting the PropBank framesets to the surgical domain. In Section 4., we describe the resulting resource highlighting that general-purpose English is significantly different from that considered in our case study – thus indicating the limited coverage, and, hence, hindered applicability, of general-purpose framebanks for surgical texts. We conclude with Section 5. summarizing our contributions and future research directions.

2. Background

In previous work different methods have been proposed to modify PropBank frames for adapting them to a specific target domain. These studies all show that the 'classic' PropBank frames can not exhaustively represent specific domains. Researchers traditionally have built NLU resources targeting general-domain English, which is syntactically and semantically different from domain-specific usage. Thus, updating the semantic frames bank can be viewed as a data-driven way for adapting algorithms for general-purpose domains to more restricted and specialized texts.

Many works on updating frames banks have been carried out in various fields, such as the clinical (Albright et al., 2013; Wang et al., 2013), the biomedical for English (Chou et al., 2006; Majewska et al., 2021) and non-English (Antony et al., 2020) texts, and other non-biomedical domains such as software analysis and cooking recipes (Jiang et al., 2020; Wang, 2015).

Wang et al. (2013) consider texts written in different laparoscopic cholecystectomy operational notes stating that the language is significantly different from general English and existing semantic resources have limited coverage of the action verbs that frequently occur in operative notes. Based on these observations, they surveyed the usage of each verb in the sample dataset to determine the verb meanings and semantic arguments of each one. In this way, they extract a set of differently used verbs, and, following the PropBank guidelines, define specific frames for them. This work, however, considers only surgical, non-robotic procedures taken from operational notes that use more schematic language than descriptions taken from textbooks for medical students and postgraduates used in our work. Furthermore, they do not consider nominalization, which is instead frequent in the robotic-surgical descriptions.

In the work from Albright et al. (2013) clinical narratives have been annotated with layers of syntactic and semantic labels to facilitate advances in clinical NLU. Following PropBank guidelines, new frames have been defined in a similar manner to the method described in our work. Although the dataset deals with a clinical language, we are considering a more specialised level, i.e., descriptions of robotic-surgical procedures, a restricted subset of the clinical domain considered by the paper (which includes for example disorders, physiology, chemicals and groups, anatomical notions, etc.). Unfortunately, the related dataset is no longer freely accessible due to copyright issues (Peng et al., 2020).

Chou et al. (2006) present a corpus of PropBank-style annotations for biomedical journal abstracts. They analyzed 30 biomedical verbs adding or modifying their meaning with respect to general-English resources. In (Majewska et al., 2021), a new resource that provides VerbNet-style (Schuler, 2006) frames for biomedical verbs was presented. The two previous works deal with a biomedical language that is still very far from the procedural surgical one.

Paper	Domain	Base	Procedural	Available
Albright et al. (2013)	Clinical narrative	PropBank	Yes*	No
Wang et al. (2013)	Clinical operative notes	PropBank	Yes	No
Chou et al. (2006)	Biomedicine (English)	PropBank	No	No
Majewska et al. (2021)	Biomedicine (English)	VerbNet	No	Yes
Jiang et al. (2020)	Recipes	PropBank	Yes	Yes
Wang (2015)	Software Requirements	PropBank	No	No
Antony et al. (2020)	Tamil (Biomedicine)	PropBank	Yes*	No

Table 1: Summary of related works. The "Dataset available" column indicates whether there exists a publicly available link to freely access the dataset. In the "Procedural" column, the value "*" means that only a part of the sentences of the whole datasets contains procedural knowledge.

Outside the medical domain, a method for automatically extracting semantic information from software requirements specifications is proposed by Wang (2015). Frequent verbs are selected from software requirement specification documents in the e-commerce domain to build the semantic frames for those verbs.

Jiang et al. (2020) instead propose a new annotated dataset for extracting information from recipes. They define ad-hoc entity types (action, food, tool, duration, temperature, condition clause, purpose clause and others) and relation types following the methodology of PropBank.

There are also several recent works presenting a Propbank with language-specific lexicons for languages other than English, such as Tamil (Antony et al., 2020) in which they perform a SRL task in Tamil Biomedicine texts, extracting domain specific verbs and related semantic roles used for the learning algorithms, Turkish (Kara et al., 2020) and Persian (Mirzaei and Moloodi, 2016). Table 1 summarizes the above works.

The presented papers show that approaching adaption of SRL algorithms by modifying semantic frames is a research direction that has proved effective in numerous fields. However, specific semantic frames for the robotic-procedural surgery domain are still missing. Our work bridges this gap by releasing the resource publicly,[1] also covering noun predicates that are frequently used in the procedural domain.

3. Building the Framebank

The overall development of the Robotic Surgery Procedural Framebank (RSPF) can be divided into two parts, namely the creation of a lexicon of frames files and the annotated corpus, following the steps described in Palmer et al. (2005). In this paper we address the first part, while the annotation of a corpus with these domain-specific frames is currently on-going.

Figure 1 shows a general overview of the domain-verbs framing process: the algorithm extracts lemmas describing actions from robotic-assisted surgical texts. How often they appear in the target domain (*freq_i*) is compared to how often they appear in OntoNotes (Weischedel et al., 2017) (*freq_i'*). If *freq_i - freq_i'* is greater than a fixed threshold, then the respective lemma is sent to a team of human linguistic experts that verify to which of the categories described in Section 3.1. the lemma belongs, modifying the corresponding frameset if necessary. The final framesets are validated by a clinician and publicly released.

Section 3.2. presents the methods we follow for extracting actionable verbs and nouns. The first one is based on keyword extraction to mainly detect actions expressed by nominalized verbs. The second one is based on Part-Of-Speech (POS) tagging to primarily detect actions expressed by verbs. Their combination offers a broad coverage of the robotic-surgery actions.

3.1. Quantifying PropBank coverage for the robotic-surgical domain

Our work represents an adaptation of the latest release (version 3.1) of PropBank (Palmer et al., 2005) to the robotic-surgical domain. We argue this adaptation is necessary for extracting meaningful robotic-surgical procedural knowledge, as the application domain is substantially different than the general English considered in PropBank. In particular, each action identified in the domain corpus (see Section 3.2.1. for details on the considered corpus) can be referred to one of the following four categories, by analyzing its semantic use in the corpus with respect to the PropBank texts:

- PRESENT: The token is already present in PropBank and there is a frame file that adequately describes the use of the predicate. For this lemma, PropBank already describes appropriate semantic roles as core entities.

- MISSING_ROLE: The token is already present in PropBank, and there is a frame file that adequately describes the use of the predicate. This frame, however, does not include domain-specific semantic roles often used in our domain. For instance, if we consider the verb *"to retract"*, PropBank offers the *"retract.01: to take back"* frame which

---

[1]To release sentences showing concrete usage examples of the domain frames, licensing arrangements have been made with the publishing houses.
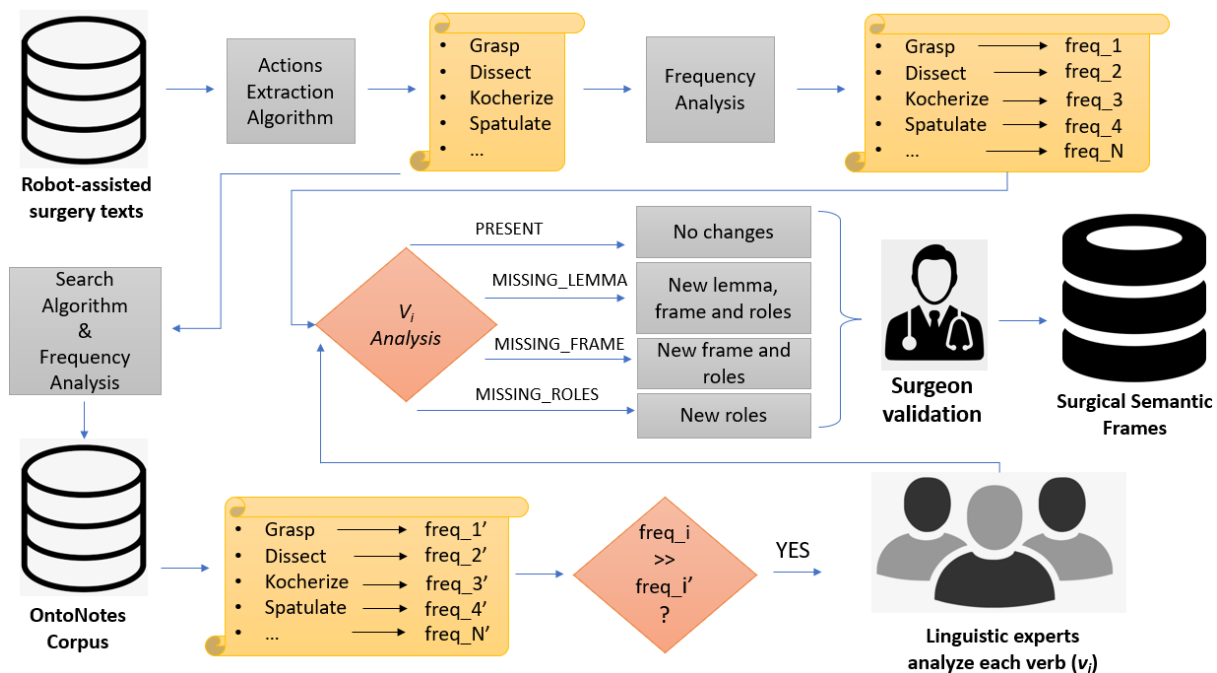
Figure 1: High level diagram of the method described in Section 3.2.3. for the framing of surgical domain verbs.

actually covers the typical meaning of the surgical domain. However, only two roles are proposed for it:

– **ARG-0:** *taker back, agent*

– **ARG-1:** *thing retracted*

The verb *to retract* is very important in our domain and used often accompanied by additional semantic roles that allow to better describe the action: the instrument used for the retraction, the technique and/or manner, and the ending point or how much to retract.

- MISSING_FRAME: The token is already present in PropBank, but a proper frame is missing, as the existing ones describe different meanings. Two different situations can arise with new instances of a lemma that already has at least one frame in PropBank: i) the new instance of the lemma is semantically entirely different from all instances that have been annotated before and there is no overlap between the old and new rolesets; ii) the frame is not completely new, but the existing one is too broad to be useful for surgery. In these cases, however, the new frame deals with a subset of the meaning captured by the old one. The lemma *approximate* in Figure 3 is an example of this second case.

An example of the first case is the verb *"to grasp"*. For it, PropBank offers a single meaning *"grasp.01: to take hold of, comprehend"* with two semantic roles:

– **ARG-0:** *grasper*

– **ARG-1:** *thing grasped*

For this lemma, in the surgical domain, we have identified a significantly different meaning: *"to clasp or embrace especially with the fingers or arms"*, with the following typical semantic roles: the grasper, the thing grasped, the instrument used for grasping, and important spatial indications for correct grasping.

- MISSING_LEMMA: The lemma is not present in PropBank. An example would be the name *"kocherization"*, which is not present in PropBank, referring to *"an operative manoeuvre to mobilize the duodenum before performing other procedures locally or before incising the duodenum"*, for which the agent and the anatomical entity typically occur as semantic roles.

To confirm the need of adapting PropBank to better support information extraction and NLU within this domain, we conducted a preliminary assessment estimating how many of the sentences of an available corpus in the robotic-surgical domain (cf. Section 3.2.1.) would be wrongly or incompletely annotated due to missing information in the original PropBank:

- ~36% of the sentences contains at least an action (with a corresponding frame in PropBank) and the mention of a domain-specific semantic role not in PropBank (MISSING_ROLE);

- ~21% of the sentences contains at least an action describing a different meaning than the

3953

Action	Ref. Verb	Modification
Placement	Place	PRESENT
Reflection	Reflect	MISSING_FRAME
Retraction	Retract	MISSING_ROLE
Exposure	Expose	PRESENT
Resection	Resect	MISSING_ROLE
Mobilization	Mobilize	MISSING_FRAME
Traction	—	MISSING_LEMMA
Administration	Administer	PRESENT
Identification	Identify	MISSING_FRAME
Excision	Excise	MISSING_ROLE

Table 2: Example of nominalized actions extracted using the method described in Section 3.2.2. with indication of the verb they refer to ("—" means missing corresponding verb).

Action	Modification
Extraperitonealize	MISSING_LEMMA
Resect	MISSING_ROLE
Spatulate	MISSING_LEMMA
Skeletonize	MISSING_LEMMA
Kocherize	MISSING_LEMMA
Insufflate	MISSING_LEMMA
Redock	MISSING_LEMMA
Detubularize	MISSING_LEMMA
Grasp	MISSING_FRAME
Incise	MISSING_ROLE

Table 3: Example domain lemmas extracted using the method described in Section 3.2.3. with indication of the type of modification required.

ones covered by its frames in PropBank (MISSING_FRAME);

- ∼2% of the sentences contains at least an action whose lemma is not in PropBank (MISSING_LEMMA);

- the remaining sentences (∼41%) could be fully annotated with information already contained in PropBank (PRESENT).

These numbers show that much less than half of the sentences in the considered dataset could be properly annotated by state-of-the-art SRL tools, unless a substantial extension of PropBank is provided.

3.2. Collecting domain-specific frames

For the extension of PropBank to the procedural robotic-surgical domain, it is first of all essential to identify those verbs (or nominalized verbs) that are typical of the surgical domain. We solved the problem using as input the datasets described in 3.2.1. and the two methods presented in 3.2.2. and 3.2.3.. For each of them, it is then necessary to check which of the categories described in 3.1. the lemma belongs to and proceed with framing as described in 3.3.. As an example, Table 2 shows 10 actions expressed by nouns identified by the method described in 3.2.2., and Table 3 shows the first 10 verbs identified by the method described in 3.2.3., with the indication of the type of modification that has been requested on PropBank.

3.2.1. Corpus used

As a corpus to extract the domain actions of the procedural robotic-surgical domain, we used the Surgical Procedural Knowledge Sentences (SPKS) dataset, presented in Bombieri et al. (2021). To the best of our knowledge, it is the only corpus dealing with this domain and has the advantage of being publicly available[2]. SPKS is a textual dataset of the surgical robotic field consisting of 1958 sentences (37022 words and 3999 unique words) manually annotated as procedural and non-procedural by an expert annotator. SPKS describes 20 different surgical procedures of four different robotic-surgery disciplines (*urology* - 39.5% of the sentences; *gastrointestinal procedures* - 30.6% of the sentences; *thoracic procedures* - 15.5% of the sentences; *gynecology* - 15.4% of the sentences).

For the comparison with general English we have instead considered the OntoNotes dataset (Weischedel et al., 2017). It is a large annotated corpus comprising various genres of text such as news, conversational telephone speech, weblogs, usenet newsgroups, broadcast and talk shows.

3.2.2. Harvesting frame-evoking nouns

In medical English, a lot of actions are expressed using nouns rather than verbs. For example, the noun *"suturation"* can be used to express an executable action as expressed by the following two semantically equivalent sentences:

- At this point, the surgeon *sutures* the vein.

- At this point a *suturation* of the vein is carried out.

In the first sentence the concept is expressed using a verb, while in the second one, the action is expressed using a nominalized verb.

For nouns, we consider the task of domain action detection as a keyword extraction problem. The keyword extraction problem is related to the identification of the lexical entities that best represent the domain. It is traditionally used in numerous fields to improve methods for browsing, indexing, topic detection, classification, contextual advertising and automatic summarizing of texts both with supervised and unsupervised approaches (Hasan and Ng, 2014). Supervised

---

approaches have promising performance in extracting keywords, but the training data requirement is a known limitation. When no annotated data are available, they must be manually generated, and this process is time consuming and subjective to human error and bias. To avoid the use of manually annotated data, unsupervised approaches have also been tested in literature, considering keyword extraction as a ranking problem.

For our purposes, we have adopted the method proposed in Campos et al. (2020). It is an unsupervised approach, built upon local text statistical features extracted from documents. It is corpus, domain, and language-independent, so it can be adapted to the robotic-surgery field, without requiring a training corpus. After a pre-processing phase, the candidate term identification step is performed, where the text is split into tokens, it is cleaned and tagged, and stopwords are identified. Then, the features extraction phase provides the input/data for computing a score for each token: a higher score indicates a more significant term.

From the output of the algorithm, we select only nominalized verbs. Since the most common morphological process involved in nominalization is derivation that can be defined as the creation of a new lexeme by the addition of an affix (i.e., a bound grammatical morphemes) (Varvara, 2017), we filter the previous results keeping only those words ending with one of the following suffixes: "-sion", "-son", "-tion", "ness", "-ment", "-ery", "-ence", "-ance", "-ure", "-ize", "-ify". False positives are finally removed from the list by manual revision.

3.2.3. Harvesting frame-evoking verbs

For verbs, we rely instead on a simple approach that compares the frequency of terms with an action role between the corpus described in 3.2.1. and a general-English corpus, OntoNotes. For each token of the domain text, we calculate its POS tag (Bird, 2006), the number of its occurrences and the sentences of the corpus in which it is used. We then retain only the tokens whose POS tag denotes a verb (i.e., *VB*, *VBP*, *VBZ*, *VBD*, *VBN* or *VBG*)[3]. We then apply lemmatization and associate each token (e.g., *"cauterized"*, *"cauterizes"*, *"cauterizing"*) to the corresponding lemma (resp., *"cauterize"*), aggregating the number of occurrences of the tokens related to the same lemma. For each obtained lemma, we then calculate the frequency with which it appears in domain sentences, and we compare it with the frequency the same lemma appears in OntoNotes. Finally, we consider as "in domain" only those lemmas that are very frequent in the domain sentences and only rarely used in OntoNotes (i.e. in which the ratio between the two frequencies is greater than a given threshold empirically set).

For example, this method identify as "of domain" verbs

like *"cauterize"*, *"detubolarize"* and *"extraperitonealize"*, because they are frequent in surgery and very rarely used in general-English and therefore the ratio between the frequencies of these verbs in the two different domains is very high. On the other hand, the method recognizes verbs such as *"need"*, *"aid"* and *"see"* as "general-English" because they appear in the two corpus with similar frequencies.

3.3. Framing of domain-actions

At the end of the processes of Sections 3.2.2. and 3.2.3., we obtain a list of domain verbs and nominalized verbs associated to a list of sentences that use them. Each lemma and the respective sentences requires a review by domain experts to understand which of the categories described in Section 3.1. it belongs to. The framing was performed by three linguistic experts with a 3-year experience in the robotic-surgical domain and validated by a clinician.

In the case MISSING_LEMMA, the lemma is not present in PropBank and thus it is an unknown word for the resource. Domain experts therefore perform the following actions: (i) addition of a new lemma to the resource; (ii) addition of a new frame to the resource for the lemma; (iii) textual definition of the meaning of that lemma in the surgical domain taken from online medical dictionaries, in particular Webster Dictionary[4] and The Free Medical Dictionary[5]; (iv) addition of appropriate semantic core-roles; (v) addition of at least one example of SRL-style annotation for the new frame. Figure 2 shows an example of the corresponding XML file for the missing lemma *"skeletonize"*.

In the case MISSING_FRAME, the lemma is already in the resource, but with inappropriate frames. In this case, domain experts perform only steps (ii)-(v). Figure 3 shows an example for the *"approximate"* lemma.

In the case MISSING_ROLE, the lemma is already in the resource with an appropriate frame, but with an inappropriate set of core-roles. In this case, steps (iv-v) are performed. Figure 4 shows an example for the *"retract"* lemma.

Finally, in the case PRESENT, the lemma is already in the resource, with an appropriate frame and core-roles. None of the previous steps are performed.

For performing step (iv), we decided to define as core an argument that either occurs with high frequency in the corpus' sentences that use that lemma (i.e., it is present in more than 50% of sentences where the lemma is used)[6] or, independently of its usage in the corpus, it is considered fundamental by domain experts for the interpretation and representation of the action.

---

[3]We only keep the verbs in this phase because the nouns describing actions have already been extracted with the method described in section 3.2.2..

[4]https://www.merriam-webster.com/medical

[5]https://medical-dictionary.thefreedictionary.com

[6]If for a lemma the associated sentences are less than 5, experts augment the dataset through web search.

```
<frameset>
  <predicate lemma="skeletonize.01">
    <roleset id="skeletonize.01"
      name="the stripping of serosal tissues from certain structures,
      e.g., the round ligament, fallopian tube, to facilitate clamping
      or excision and minimize bleeding.">
      <note>Frames file for 'skeletonize' based on a definition given by
        "Concise Dictionary of Modern Medicine"</note>
      <roles>
        <role descr="agent" f="" n="0"> </role>
        <role descr="anatomical part skeletonized" f="" n="1"> </role>
      </roles>
      <example name="skeletonize - surgery" src="" type="">
        <text>The uterine vessels are then carefully skeletonized.</text>
        <arg f="" n="1">The uterine vessels</arg>
        <arg f="MNR" n="m">carefully</arg>
        <rel f="">skeletonized</rel>
      </example>
    </roleset>
  </predicate>
</frameset>
```
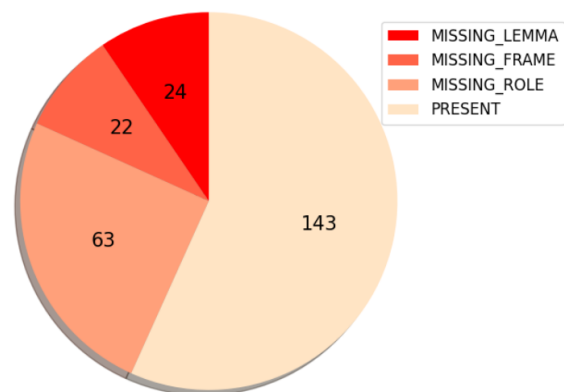
Figure 2: XML file for the "skeletonize" lemma.

```
<roleset id="approximate.02" name="To come near in position, to bring near.">
  <note>Frames file for 'approximate' based on survey of sentences in
    the SPKS corpus.</note>
  <roles>
    <role descr="agent" f="" n="0"></role>
    <role descr="entity coming close" f="" n="1"></role>
    <role descr="entity coming close to" f="" n="2"></role>
    <role descr="instrument" f="" n="3"></role>
    <role descr="other useful spatial indications" f="" n="4"></role>
  </roles>
  <example name="approximate - surgery" src="" type="">
    <text>Next, the musculofascial plate of the rectourethralis is
      approximated to the posterior neobladder, approximately 2-cm
      posterior to the planned urethral aperture, using 3-0
      polyglactin suture.</text>
    <arg f="tmp" n="m">Next</arg>
    <arg f="" n="1">the musculofascial plate of the rectourethralis</arg>
    <rel f="">approximated</rel>
    <arg f="" n="2">to the posterior neobladder</arg>
    <arg f="" n="4">approximately 2-cm posterior to the planned urethral
      aperture</arg>
    <arg f="" n="3">using 3-0 polyglactin suture</arg>
  </example>
</roleset>
```

Figure 3: XML file for the "approximate" lemma.

```
<frameset>
  <predicate lemma="retract">
    <roleset id="retract.01" name="to take back.">
      <aliases>[...]</aliases>
      <roles>
        <role descr="taker back, agent" f="PAG" n="0">
          <vnrole vncls="10.1" vntheta="Agent"/>
        </role>
        <role descr="thing retracted" f="PPT" n="1">
          <vnrole vncls="10.1" vntheta="Theme"/>
        </role>
        <!--ADDED-->
        <role descr="instrument used for retraction" f="" n="2">
        </role>
        <role descr="technique / manner" f="" n="3"></role>
        <role descr="ending point or how much" f="" n="4"></role>
        <!--*****-->
      </roles>
      [...]
      <!--ADDED-->
      [...]
      <example name="Surgery with arg2 and arg4" src="" type="">
        <text>The mediastinal pleura is then retracted posteriorly
          with cadiere forceps (arm 3).</text>
        <arg f="" n="1">The mediastinal pleura</arg>
        <arg f="tmp" n="m">then</arg>
        <rel f="">retracted</rel>
        <arg f="" n="4">posteriorly</arg>
        <arg f="" n="2">with cadiere forceps (arm 3)</arg>
      </example>
      <example name="Surgery with arg3" src="" type="">
        <text>The left lobe of the liver can easily be retracted
          using the suture-gauze liver suspension method.</text>
        <arg f="" n="1">The left lobe of the liver</arg>
        <rel f="">retracted</rel>
        <arg f="" n="3">using the suture-gauze liver suspension
          method.</arg>
      </example>
      [...]
      <!--*****-->
    </roleset>
  </predicate>
</frameset>
```

Figure 4: XML file for the "retract" lemma.

Figure 5: Changes made to the original resource

The framing phase is very expensive because it is carried out manually by personnel who must have expertise both in linguistics (SRL annotations in PropBank style) and in the robotic-surgical domain. It took the domain experts about 80 hours to complete the task.

4. The Robotic Surgery Procedural Framebank

Using the method described in Section 3. we have analyzed 252 lemmas. At least one modification among those described in Section 3.1. have been requested in 109 cases. In particular, we have added 24 new lemmas (verbs or nouns) that describe very specific actions of the surgical domain. We have added 22 new frames to existing lemmas describing meanings not already covered by PropBank. Considering both the new lemmas added, new frames added to existing lemmas and the new role-cores added to existing frames, we have added a total of 244 core-roles.The type of changes is summarized in Figure 5, where the frequencies of different modification introduced are reported.

Figure 6 shows the semantic type of core-roles we added for the analysed lemmas. The figure considers all core-roles added, both in existing frames and in new frames. While the number of ARG-0 and ARG-1 inserted is quite high (class "A"), they all derive from new frames not yet present in the standard PropBank. In class "C" the group *other sp. attributes* ("other spatial attributes") refers to other general spatial information, e.g., *orientation*, *spatial constraints* or other recurring frame-specific spatial attributes.

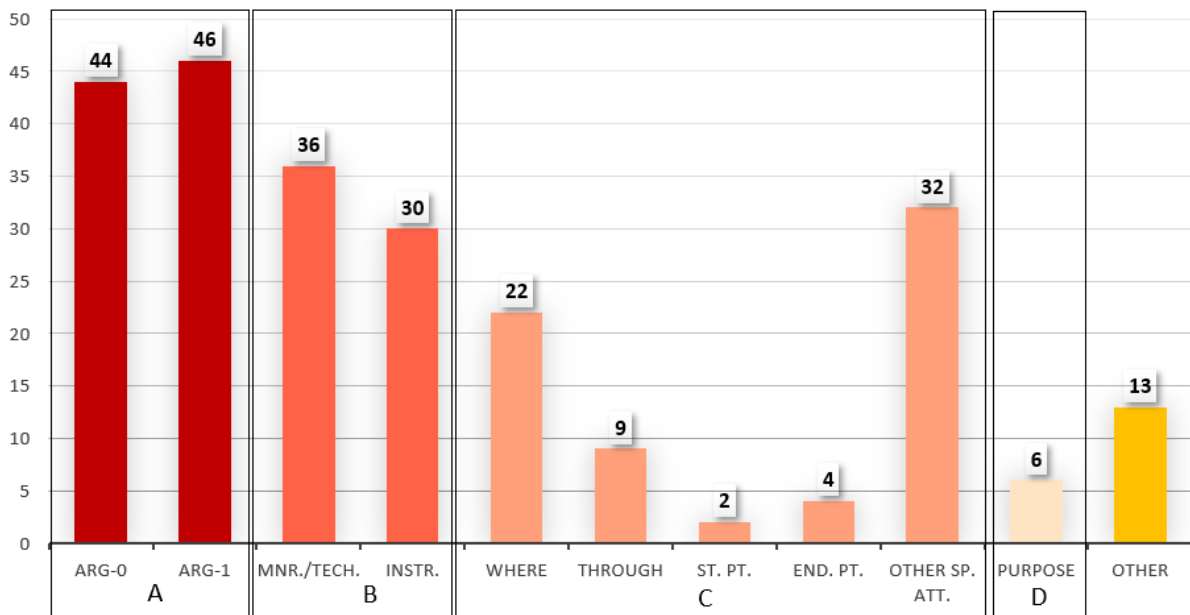The nature of the semantic roles inserted highlights

Figure 6: Changes made to the original resource. The class labelled with letter "A" is composed by ARG-0 (proto-agent: who performs an action) and the ARG-1 (proto-patients: who undergoes the action). The class named with letter "B" contains indeed the instrument used to carry out an action and the surgical technique to follow to properly execute it. The class named with letter "C" contains all those roles that are related with the concept of space. The class "D" deals with the purpose of a particular action. Finally we leave in the "Other" class all those roles that are very specific to a particular lemma and does not fit in any of the above classes

that in the surgical procedural language it is of utmost importance to indicate, for each action that describes an operation, who performs the action (ARG-0), the anatomical part that undergoes the action (often ARG-1), the instrument with which to perform the action, the surgical technique to adopt, the purpose, and a series of spatial information that helps locate the target anatomy within the human body.

Overall, the distribution of newly introduced and modified lemmas and frames indicates that the extension of PropBank to cover the robotic-surgical domain is substantial and provides a solid basis for the annotation of the SPKS corpus.

5. Conclusions

In this paper we presented a PropBank extension for the robotic-surgical domain. To this end, we have adopted two methods to extract procedural robotic-surgical domain nouns and verbs: one based on keyword extraction, and another one based on frequency comparison between lemmas used in procedural robotic-surgery and in general-English corpus. The resulting resource highlights that the lemmas used in procedural robotics-surgical texts are substantially different from those found in general-purpose English. Three human linguistic experts carried out the frame annotation task. One clinician finally validated the results. We make our resource publicly available to foster further research on

the topic of domain adaptation for SRL in the healthcare, and specifically surgical, domain.

Building upon previous work, we envision extracting procedural knowledge from this challenging domain on the basis of SRL techniques. Thus, our plan is to use this novel resource for annotating domain-specific texts from SPKS with semantic role information, so as to benchmark the coverage and quality of state-of-the-art systems trained on standard newswire text (i.e., 'vanilla' PropBank), as well as to investigate the potential benefits of training/fine-tuning using domain-specific annotations.

Downloads

The resulting PropBank extension for the robotic-surgical domain is made available at: `https://gitlab.com/altairLab/robotic-surgery-propositional-bank`.

Acknowledgements

6. Bibliographical References

Albright, D., Lanfranchi, A., Fredriksen, A., IV, W. F. S., Warner, C., Hwang, J. D., Choi, J. D., Dligach, D., Nielsen, R. D., Martin, J. H., Ward, W. H., Palmer, M., and Savova, G. K. (2013). Towards comprehensive syntactic and semantic annotations of the clinical narrative. *J. Am. Medical Informatics Assoc.*, 20(5):922–930.

Antony, J. B., Paul, N. R. R., and Mahalakshmi, G. S. (2020). Entity and verb semantic role labelling for tamil biomedicine. In Purushothama B. R., et al., editors, *Mining Intelligence and Knowledge Exploration*, pages 72–83, Cham. Springer International Publishing.

Baker, C. F., Fillmore, C. J., and Lowe, J. B. (1998). The berkeley framenet project. In Christian Boitet et al., editors, *36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, COLING-ACL '98, August 10-14, 1998, Université de Montréal, Montréal, Quebec, Canada. Proceedings of the Conference*, pages 86–90. Morgan Kaufmann Publishers / ACL.

Bird, S. (2006). NLTK: The Natural Language Toolkit. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, pages 69–72, Sydney, Australia, July. Association for Computational Linguistics.

Bombieri, M., Dall'Alba, D., Ramesh, S., Menegozzo, G., Schneider, C., and Fiorini, P. (2020). Jointsspace metrics for automatic robotic surgical gestures classification. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3061–3066.

Bombieri, M., Rospocher, M., Dall'Alba, D., and Fiorini, P. (2021). Automatic detection of procedural knowledge in robotic-assisted surgical texts. *International journal of computer assisted radiology and surgery*, 16, 04.

Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., and Jatowt, A. (2020). Yake! keyword extraction from single documents using multiple local features. *Information Sciences*, 509:257–289.

Chen, L., Jiang, Z., Xiao, J., and Liu, W. (2021). Human-like controllable image captioning with verb-specific semantic roles. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 16846–16856. Computer Vision Foundation / IEEE.

Chou, W.-C., Tsai, R. T.-H., Su, Y.-S., Ku, W., Sung, T.-Y., and Hsu, W.-L. (2006). A semi-automatic method for annotating a biomedical Proposition Bank. In *Proceedings of the Workshop on Frontiers in Linguistically Annotated Corpora 2006*, pages 5–12, Sydney, Australia, July. Association for Computational Linguistics.

Demner-Fushman, D., Chapman, W. W., and McDonald, C. J. (2009). What can natural language processing do for clinical decision support? *Journal of biomedical informatics*, 42 5:760–72.

Do, Q. T. N., Bethard, S., and Moens, M.-F. (2015). Domain adaptation in semantic role labeling using a neural language model and linguistic resources. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(11):1812–1823.

Gildea, D. and Jurafsky, D. (2002). Automatic labeling of semantic roles. *Comput. Linguistics*, 28(3):245–288.

Hasan, K. S. and Ng, V. (2014). Automatic keyphrase extraction: A survey of the state of the art. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1262–1273, Baltimore, Maryland, June. Association for Computational Linguistics.

Humphreys, L., Boella, G., van der Torre, L., Robaldo, L., Caro, L. D., Ghanavati, S., and Muthuri, R. (2021). Populating legal ontologies using semantic role labeling. *Artif. Intell. Law*, 29(2):171–211.

Jiang, Y., Zaporojets, K., Deleu, J., Demeester, T., and Develder, C. (2020). Recipe instruction semantics corpus (risec): Resolving semantic structure and zero anaphora in recipes. In Kam-Fai Wong, et al., editors, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing, AACL/IJCNLP 2020, Suzhou, China, December 4-7, 2020*, pages 821–826. Association for Computational Linguistics.

Kara, N., Aslan, D. B., Marsan, B., Bakay, Ö., Ak, K., and Yildiz, O. T. (2020). Tropbank: Turkish propbank v2.0. In *LREC*.

Majewska, O., Collins, C., Baker, S., Björne, J., Brown, S. W., Korhonen, A., and Palmer, M. (2021). Bioverbnet: a large semantic-syntactic classification of verbs in biomedicine. *J. Biomed. Semant.*, 12(1):12.

Mirzaei, A. and Moloodi, A. (2016). Persian proposition bank. In *LREC*.

Nyamulani, N. and Mulwafu, W. (2018). The quality of hand-written operative notes in a surgical unit at queen elizabeth central hospital (qech), malawi: A prospective completed audit loop study. *Malawi Medical Journal*, 30:86 – 89.

Palmer, M., Kingsbury, P. R., and Gildea, D. (2005). The proposition bank: An annotated corpus of semantic roles. *Comput. Linguistics*, 31(1):71–106.

Peng, Y., Zhang, Z., Wang, X., Yang, L., and Lu, L. (2020). Chapter 5 - text mining and deep learning for disease classification. In S. Kevin Zhou, et al., editors, *Handbook of Medical Image Computing and Computer Assisted Intervention*, The Elsevier and MICCAI Society Book Series, pages 109–135. Academic Press.

Rebmann, A. and van der Aa, H. (2021). Extracting

semantic process information from the natural language in event logs. In Marcello La Rosa, et al., editors, *Advanced Information Systems Engineering*, pages 57–74, Cham. Springer International Publishing.

Sadhu, A., Gupta, T., Yatskar, M., Nevatia, R., and Kembhavi, A. (2021). Visual semantic role labeling for video understanding. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 5589–5600. Computer Vision Foundation / IEEE.

Schuler, K. K. (2006). *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. Ph.D. thesis, University of Pennsylvania.

Varvara, R. (2017). *Verbs as nouns: empirical investigations on event-denoting nominalizations*. Ph.D. thesis, University of Trento, 05.

Wang, Y., Pakhomov, S., and Melton, G. (2013). Predicate argument structure frames for modeling information in operative notes. In *MEDINFO 2013 - Proceedings of the 14th World Congress on Medical and Health Informatics*, number 1-2 in Studies in Health Technology and Informatics, pages 783–787. IOS Press. 14th World Congress on Medical and Health Informatics, MEDINFO 2013 ; Conference date: 20-08-2013 Through 23-08-2013.

Wang, Y. (2015). Semantic information extraction for software requirements using semantic role labeling. In *2015 IEEE International Conference on Progress in Informatics and Computing (PIC)*, pages 332–337.

Weischedel, R. M., Hovy, E. H., Marcus, M. P., and Palmer, M. (2017). Ontonotes : A large training corpus for enhanced processing. In *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation*. Springer.

Xu, K., Wu, H., Song, L., Zhang, H., Song, L., and Yu, D. (2021). Conversational semantic role labeling. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29:2465–2475.