Reading as an Enabling Technology: Informing Surgical Robots with Spatial Information

Maria-Camilla Fiazza[1], Marco Bombieri[1], and Paolo Fiorini[1]

[1]*Department of Computer Science, University of Verona, Italy*
Contact: mariacamilla.fiazza@univr.it

INTRODUCTION

Reference surgical textbooks are a source of high-quality information that has been compiled by experts and heavily vetted. This information is not easily available elsewhere. Making it available to robotic surgical systems requires that engineers extensively consult with medical experts while bridging a wide disciplinary gap.

Extracting high-level information from surgical textbooks *automatically* can enable multiple forms of progress in the development of autonomous systems for surgery. Firstly, it can reduce system development costs, redirecting the tool of expert consultation toward the more difficult questions about context. Secondly, it can enable autonomous planning of surgical procedures, and contribute significantly to the situation awareness and usefulness of devices that operate in assistive mode. Thirdly, this new technology can enable on-demand availability of supplemental information—perhaps intra-operatively in response to unexpected events.

The two key challenges are extracting structured high-level information from texts written in natural language and presenting the information thus extracted to the systems modules that need it, in a format that is suitable for use. This contribution focuses on the latter challenge, a key integration step toward cognitive surgical robotics.

Semantic Content and Navigation

The semantic content of each of the sentences in a surgical textbook is varied and of diverse nature. Ontological information (information about the properties and relationships of objects in the scene) is of interest for modules that perform semantic segmentation or scene understanding [1]. High-level planners rely, instead, on procedural knowledge, which leverages the ontology to describe who carries out an action, when and how. The key architectural question is which modules in a robotic system are responsible for integrating which type of information (and when), so that sentences in natural language can be mapped at the end of the process into an executable representation, where parameters have known values.

Of critical importance is the ability to decode a type of information that is central in all surgical texts—*spatial* information—and to enable a form of spatial reasoning. Spatial information can be encoded at word level, in addition to phrase or clause level, and can be carried by verbs (retract, lift), adjectives (anterior, renal), prepositions (through, via, in) and a number of morphological affixes (retro-, trans-). In surgical robotics, the environment is the patient's *body*. Anatomical objects (e.g., organs) carry spatial information both directly, by encoding their location and spatial extension, and indirectly, serving as terms of reference, or as navigational landmarks.

We explored the feasibility of this translation task using partial nephrectomy as a case study. In-depth analysis of the information carried by a set of representative sentences yielded architectural requirements and a strategy to approach the larger work of automatic conversion. The result is a set of word-level transformations and a minimal working architecture informed by the requirements.

MATERIALS AND METHODS

From a database of 1958 sentences from surgical textbooks on robot-assisted interventions, annotated as procedural or not [2], we selected all the sentences related to partial nephrectomy, our case-study procedure, yielding the 44-sentence description from [3]. Filtering out the non-procedural sentences and mapping the syntactic elements onto general semantic roles (semantic parsing or logical analysis) can be accomplished with state-of-the-art tools. We used the methods in [2] to perform the filtering; we leveraged semantic banks [4] to disambiguate text and yield representations such as those in Step 2 in Fig. 1, where phrases are assigned to roles such as subject, object, instrument, time, etc.

In-depth manual analysis of semantic content resulted in defining a set of transformations to translate spatial information contained in the high-level roles of Step 2 (Fig. 1). These transformations can be used to automate the procedure on new text. Phrases such as "below the instrument" are resolved at run-time and expressed in the form of a procedure: query the system (namebank) for the object `Instr` whose label is the string "instrument"; query the system (semantic segmentation module) for the volume occupied by `Instr`; apply the fuzzy operator `below` to the volume. `below` is configured to select for the absolute negative z direction if no other axis is specified (the patient may not be fully horizontal). Spatial transformations are represented by fuzzy sets and dilation operators, as described, for example in [5]. Depending on the intended use, all spatial information is labeled as one of: {location, configuration, trajectory, constraint}.

Translation is word-based, because the natural-language semantic is as well. For example, *lift* implicitly refers to a global reference axis pointing upward, whereas *rotate*

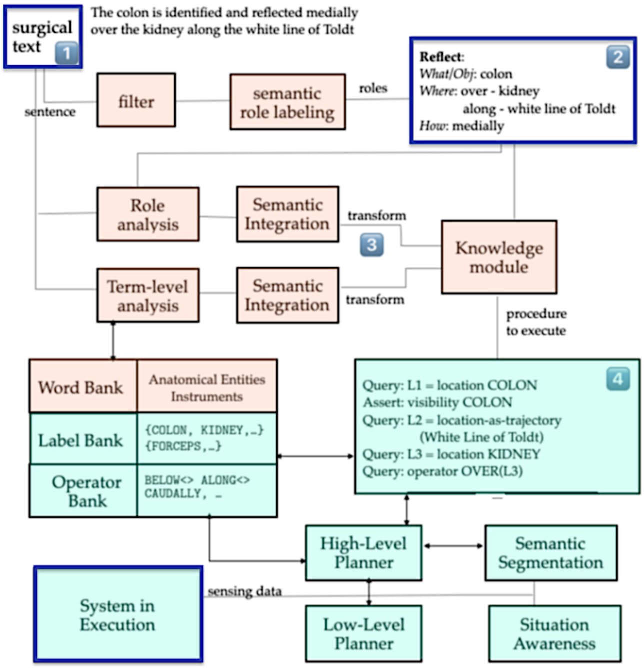

Fig. 1. Processing architecture to resolve domain-specific information in natural language and make it available to a robotic system. Orange: textual translation system; Teal: surgical robotic system; Blue border: forms of procedural information. Normal lines represent input-to relationships; arrows represent active communication (querying) between components.

implicitly refers to a local axis specific to the organ. The knowledge module integrates all spatial information that is available "at compile time" in the form of operators that augment the data structure at Step 2; using the transforms at Step 3, it generates the program in Step 4.

The knowledge module informs system components by passing along hybrid data structures, carrying information from multiple levels of abstraction and composed of a mix of semantic information that has already been resolved and information that is yet to be resolved. The latter is represented as a sequence of instructions that include system queries. The overall system architecture must be able to serve such queries.

RESULTS AND DISCUSSION

We identified in the text the following elements: words to reference organs (e.g. colon, kidney/renal); names for tissue types that can be found as *part* of organs or that qualify the parts (e.g., fat, tumor); names for landmarks (accesses, ports, and organ-level landmarks such as the hilum); words for spatial concepts (e.g., surface, working space); spatial operators or modifiers (e.g., via, retro-, along); verbs that involve spatial concepts.

A state-of-the-art semantic labeling tool [6] was able to correctly segment units of spatial information governed by modifiers, verbs or nominalized verbs 77.8% of the time. However, careful inspection of the automated output in Step 2 showed that only 61.5% of the discoverable spatial labels were detected. Although powerful, available semantic banks do not perform well on texts as domain-dependent and semantically complex as surgical texts. Research is needed to adapt semantic data banks to the surgical domain and to enable selection of the correct text-to-procedure transformation (Step 3), so as to automate the translation process to the extent possible.

In Fig. 1, the entire system contributes to decoding the spatial information into an executable form. Not all information *can* be provided, even in principle, because textbooks do not specify by how much to lift an anatomical structure, or what the stopping conditions are. What is still left unspecified is suitable to being learned in a data-driven manner, especially via learn-by-example paradigms. It is possible in this manner to constrain the system's learning to semantically meaningful units, which is helpful when attempting to offer safety guarantees on system performance, as suggested in [7].

CONCLUSIONS

In this contribution, we reported on ongoing preliminary work to simultaneously design a set of (word-level and phrase-level) semantic mappings and a system architecture, to support the progressive integration of semantic information. Our approach can also contribute missing word-meaning pairs to the current state-of-the-art semantic data banks. This is enabling technology for a long-term goal. Intermediate steps consist in increasing the level of autonomy of surgical devices, especially in assistive mode. Once mature, the technology can be extended to instructions given by competent physician in natural language directly to the robot.

Our next steps consist in addressing the contextual disambiguation problem for select operators (`anterior/ly` and `along`) and in verifying how well our strategy for translation generalizes to texts by other authors describing the same robot-assisted procedure.

ACKNOWLEDGEMENTS

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 742671 "ARS").

REFERENCES

[1] A. O. Alarcos, D. Beßler, A. M. Khamis *et al.*, "A review and comparison of ontology-based approaches to robot autonomy," *Knowledge Engineering Review*, vol. 35, p. e5, 2020.

[2] M. Bombieri, M. Rospocher, D. Dall'Alba, and P. Fiorini, "Automatic detection of procedural knowledge in robotic-assisted surgical texts," *International Journal of Computer-Assisted Radiology and Surgery*, vol. 16, 04 2021.

[3] J. S. Parihar and C. Lau, *Robot-Assisted Partial Nephrectomy*. Cham: Springer International Publishing, 2018, pp. 103–111.

[4] M. Palmer, D. Gildea, and P. Kingsbury, "The Proposition Bank: An Annotated Corpus of Semantic Roles," *Computational Linguistics*, vol. 31, no. 1, pp. 71–106, 03 2005. [Online]. Available: https://doi.org/10.1162/0891201053630264

[5] I. Bloch, "Mathematical Morphology and Spatial Reasoning: Fuzzy and Bipolar Setting," *TWMS Journal of Pure and Applied Mathematics*, vol. 12, no. 1, pp. 104–125, 2021.

[6] Allen Institute for AI, "AllenNLP Demo," Last access: 2022/04/01. [Online]. Available: https://demo.allennlp.org/semantic-role-labeling

[7] M.-C. Fiazza and P. Fiorini, "Design for Interpretability: Meeting the Certification Challenge for Surgical Robots," in *2021 IEEE International Conference on Intelligence and Safety for Robotics (ISR)*, 2021, pp. 264–267.