

PatientNarr: Towards generating patient-centric summaries of hospital stays

Barbara Di Eugenio, Andrew D. Boyd, Camillo Lugaresi, Abhinaya Balasubramanian, Gail M. Keenan, Mike Burton, Tamara G. Rezende Macieira, Karen Dunn Lopez

University of Illinois at Chicago
Chicago, IL, USA

Carol Friedman
Columbia University
New York, NY, USA

Jianrong Li, Yves A. Lussier
University of Arizona
Tucson, AZ, USA

Abstract

PatientNarr summarizes information taken from textual discharge notes written by physicians, and structured nursing documentation. It builds a graph that highlights the relationships between the two types of documentation; and extracts information from the graph for content planning. SimpleNLG is used for surface realization.

1 Introduction

Every year, 7.9% of the US population is hospitalized (CDC, 2011). Patients need to understand what happened to them in the hospital, and what they should do after discharge. *PatientNarr* will ultimately be able to generate concise, lay-language summaries of hospital stays. We hypothesize that such summaries will help patients take care of themselves after they are discharged, and supplement current approaches to patient education, which is not always effective (Olson and Windish, 2010).

PatientNarr needs to summarize documentation that is currently **segregated by profession**; as a minimum, as *physician discharge notes* and as *nursing plans-of-care*. We contend that both sources are necessary to provide the patient with full understanding, also because much of the direct care provided by nurses will need to be continued following discharge (Cain et al., 2012).

In our case, PatientNarr summarizes data that is heterogeneous (textual for physician discharge notes, structured for nursing documentation). This paper describes the steps we have undertaken so far: (a) To demonstrate that physician and nurse documentations diverge, we map both to a graph, and study the relationships therein. This graph supports content planning. (b) We have developed the pipeline that extracts the information to

be communicated, and renders it in English via SimpleNLG (Gatt and Reiter, 2009).

Related work. NLG and Summarization in the biomedical domain have been pursued for a few years (Di Eugenio and Green, 2010), but most work addresses health care personnel: to navigate cancer patients' medical histories (Hallett, 2008; Scott et al., 2013); to generate textual summaries describing a hospitalized infant for nurses (Portet et al., 2009); to generate reports of care for hand-off between emergency workers (Schneider et al., 2013). Most applications of NLG that target patients focus on behavioral changes (Reiter et al., 2003), or patient counseling (Green et al., 2011). Only few NLG systems attempt at generating personalized medical information from medical records or data (Williams et al., 2007; Mahamood and Reiter, 2011).

2 A motivating example

So far, we have gained access to 28 de-identified discharge notes of cardiology patients from the University of Illinois Hospital and Health Science System (UIHSS). Figure 1 shows about 20% of the physician discharge notes for Patient 9. It is difficult to understand, not only because of jargon, but also because of ignorance of relevant domain relationships. Importantly, these notes do not talk about issues that are potentially important for the patient, like his state of mind, which are more often addressed by nurses. In our case, the nursing documentation is not textual, but entered via the HANDS tool, and stored in a relational database (Keenan et al., 2002). A tiny portion of the initial plan-of-care (POC) for Patient 9 is shown in Figure 2 (this nursing data is reconstructed, see Section 3). One POC is documented at every formal handoff (admission, shift change, or discharge). HANDS employs the NANDA-I taxon-

Patient was admitted to Cardiology for new onset a fib in RVR. Was given an additional dose of diltazem 30mg po when first seen. Patient was started on a heparin drip for possible TEE and cardioversion. Overnight his HR was in the 100-110s; however did increase to 160s for which patient was given 120mg er of diltazem. HR improved; however, in the morning while awake and moving around, HR did increase to the 130-140s. Patient was given another dose of IV dilt 20mg. [...] Upon discharge was given two prescriptions for BP, HCTZ and losartan given LVH seen on echo. Patient was counseled on the risks of stroke and different options for anticoagulation. [...]

Figure 1: Excerpt from physician discharge notes (Patient 9)

omy of nursing diagnoses, represented by squares in Figure 2; the Nursing Outcomes Classification (NOC) – circles; and the Nursing Interventions Classification (NIC) – triangles (NNN, 2014). In Figure 2, *Acute Pain* is a diagnosis, and *Anxiety Level* and *Pain Level* (some of) its associated outcomes. *Anxiety Reduction* is an intervention associated with *Anxiety Level*; *Pain Management* and *Analgesic Administration* are interventions associated with *Pain Level*. A scale from 1 to 5 indicates the initial value associated with an outcome (i.e., the state the patient was in when s/he was admitted), the expected rating, and the actual rating at discharge. In Figure 2, the current level for *Pain Level* and *Anxiety Level* is 2 each, with an expected level of 5 at discharge, i.e., no pain/anxiety.

Patient	Patient 9	Medic DX	Atrial fibrillation with rapid ventricular response
DOB	MM/DD/YY	Allergies	NKA
POC	6/20/2013	Physician	XX
Shift	7:00a - 7:00p	Current Rating	Expected Rating
	Acute Pain		
	Anxiety Level	2	(5)
	Anxiety Reduction		
	Pain Level	2	(5)
	Pain Management		
	Analgesic Administration		

Figure 2: Excerpt from nursing documentation

Figures 1 and 2 suggest that physician and nursing documentations provide different perspectives on patients: e.g., *Anxiety* is not even mentioned in the discharge notes. One of the authors (a nursing student) wrote summaries for five of the 28 discharge summaries and their corresponding HANDS POCs – Figure 3 shows the summary for Patient 9. This initial round of human authoring was meant to provide some preliminary guidelines to generate automatic summaries. Please see Section 5 for our plans on obtaining a much larger quantity of more informed human-authored sum-

maries.

3 Extracting relationships between physician notes and nursing data

To extract and relate information from our two sources, we rely on UMLS, MedLEE and HANDS. UMLS, the Unified Medical Language System (NLM, 2009), includes 2.6 million concepts (identified by Concept Unique Identifiers or CUIs) organized in a network. Importantly, many different medical and nursing terminologies have been incorporated into UMLS, including those used by HANDS (NANDA-I, NIC and NOC). UMLS provides mapping between their concepts and CUIs, via 8.6 million concept names and relationships between terminologies. Some relationships are of a hierarchical nature, where one concept is narrower than the other (e.g., Chest X-ray and Diagnostic radiologic examination).

MedLEE is a medical information extraction system (Friedman et al., 2004). In its semi-structured output, recognized entities are mapped to the corresponding CUI in UMLS.

HANDS has not been adopted at UIHSS yet. Hence, we reconstructed HANDS POCs for those 28 patients on the basis of 40,661 cases collected at four hospitals where HANDS is in use. For each of the 28 patients, the same nursing student who authored the five summaries, selected similar cases, and used them to produce high-quality records consistent with actual nursing practice.

To relate physician and nursing documentations, we seed a graph with two sets of CUIs: those returned by MedLEE as a result of processing the physician discharge notes; and the CUIs corresponding to all the NANDA-I, NIC and NOC terms from the HANDS POCs. We then grow the graph by querying UMLS for the set of concepts related to each of the concepts in our set; the concepts that were not already part of the graph are then used to begin a new round of growth (we stop at round 2, to keep the time used by UMLS to answer, reasonable). From this graph, we keep the concepts that either belong to one of the

You were admitted with new onset of atrial fibrillation. You reported feeling weakness, chest pressure and increased shortness of breath. You reported acute pain and you were anxious. During your hospitalization you were treated with analgesics for your pain and pain management was performed by the nursing team. Your shortness of breath improved. Your decreased cardiac output was treated with medication administration and knowledge about cardiac precautions, hypertension management, your treatment regimen and the prescribed medication were taught to you by the nurses. A Transophageal Echocardiography was performed. You met the expected outcomes for your condition and you were discharged under the condition improved for your home. You have an appointment scheduled at Union Medical Center on [DATE] with Physician [DR.]. The list of your medications is attached to this discharge.

Figure 3: Human-authored summary (Patient 9)

source lists, or that are required to form a connection between a doctor-originated concept and a nurse-originated concept that would otherwise remain unconnected. All other concepts are removed. The result is a graph with several separate connected components, which correspond to clusters of related concepts occurring in the discharge notes or in the plans of care, or forming connections between the two sources.

We count distances in terms of relationships traversed, starting from the nursing concepts since they are fewer, and since path traversal is reversible.¹ Concepts can overlap; or be directly connected (distance one); or be directly connected through an intermediate concept (distance two). We do not consider distances beyond two. Table 1 shows results for our specific example, Patient 9, and average results across our 28 test cases. As we can see, there are very few concepts in common, or even at distance 1. Our results provide quantitative evidence for the hypothesis that physicians and nurses talk differently, not just as far as terminology is concerned, but as regards aspects of patient care. This provides strong evidence for our hypothesis that a hospitalization summary should include both perspectives.

4 Automatically generating the summary

In this baseline version of PatientNarr, we focused on understanding how the vital parameters have improved over time, the treatment given for improvement and the issues addressed during the process. The summary generated by PatientNarr for Patient 9 is shown in Figure 4. We extract information of interest from the graph obtained at the previous step; we couch it as features of phrasal constituents via the operations provided by the SimpleNLG API. SimpleNLG then assembles grammatical phrases in the right order, and helps

¹In UMLS, any relationship from concept A to concept B, has a corresponding relationship from B to A, not necessarily symmetric.

in aggregating related sentences.

Since there are far fewer nursing than doctor concepts, we start from the NANDA-I codes, i.e., the diagnoses. The name associated in UMLS with the corresponding CUI is used. For each NANDA-I node, we highlight the treatments given (the NIC codes), e.g. see the sentence starting with *Acute onset pain was treated [...]* in Figure 4. For both diagnosis and treatments, we attempt to relate them to doctor’s nodes. Specifically, we exploit the relationships in the UMLS ontology and include nodes in the graph we constructed that are at distance 1 or 2, and that are either doctor’s nodes, or intermediate nodes that connect to a doctor’s node. For example, in *Dysrhythmia management is remedy for tachycardia and Atrial Fibrillation*, *Dysrhythmia management* is a NIC intervention that is related to *Cardiac Arrhythmia*; in turn, *Cardiac Arrhythmia* is a direct hypernym of *tachycardia* and *Atrial Fibrillation* which were both extracted from the physician notes by MedLEE. *Cardiac Arrhythmia* was discovered by our graph building procedure, as described in Section 3.

We then highlight what improved during the hospital stay. As we mentioned earlier, the NOC codes (outcomes) are associated with a scale from 1 to 5 which indicates the initial value, the expected rating, and the actual rating at discharge. If the relative increase between admission and discharge encompasses more than 2 points on the scale, it is considered significant; if it encompasses 1 or 2 points, it is considered slight. In those cases in which more than one outcome is associated with a diagnosis, but improvement is not uniform, we include a cue “On the other hand”. For Patient 9, in the last POC recorded just before discharge, *Anxiety Level* is up 2 points (to 4), whereas *Pain Level* is up 3. We also indicate to the patient if the final rating reached the rating that was initially hoped for; it did not for *Anxiety Level* (See the two sentences starting from *Pain level and Vi-*

	# CUIs from Discharge Notes	# CUIs from Nursing POCs	# of common CUIs	# of CUI pairs at Distance 1	# of CUI pairs at Distance 2
Patient 9	83.00	28.00	0.00	3.00	13.00
Average	90.64	22.43	0.46	3.00	9.11

Table 1: Concept overlap in discharge notes and nursing POCs

You were admitted for atrial fibrillation with rapid ventricular response. Acute onset pain related to pain was treated with pain management, monitoring of blood pressure, temperature, pulse rate and respiratory rate and administration of analgesic. Pain level and vital signs have improved significantly and outcomes have met the expectations. On the other hand, level of anxiety has improved slightly. Cardiac Disease Self-Management, Disease Process (Heart disease), Hypertension Management, Cardiac Disease Management, Hypertension Management and Treatment Regimen were taught. Low Cardiac Output related to heart failure was treated with cardiac precautions, monitoring of blood pressure, temperature, pulse rate and respiratory rate, and dysrhythmia management. Dysrhythmia management is remedy for tachycardia and atrial fibrillation. As a result, cardiac pump effectiveness, cardiopulmonary status and cardiac tissue perfusion status have improved slightly. Actual Negative Breathing Pattern related to respiration disorders was treated with respiration monitoring. Respiratory Status has improved significantly. You have an appointment at Union Medical Center on DATE at TIME. The list of medication is attached to this discharge.

Figure 4: PatientNarr generated summary (Patient 9)

tal signs [...]. On the other hand, [...]. For the moment, we do not mention outcomes for which no improvement, or a setback, has been recorded.

The summary also includes: mentions of education that has been imparted; and reminders of future appointments and of medicines to be taken.

5 Future Work

The research described in this paper lays the foundations for our project, but clearly much work remains to be done. To start with, we plan to build a corpus of gold-standard summaries, in order to derive (semi-)automatic models of the information to be included from physician notes and from nursing documentation. The five human authored summaries we currently have at our disposal are not sufficient, neither in quality nor (obviously) in quantity. We intend to inform their generation via a number of focus groups with all stakeholders involved: patients, doctors and nurses. To start with, the five summaries we do have were presented to the Patient Advisory Board of an unrelated project. These two patients noted that all unfamiliar terms should be explained, and that what the patient should do to improve their health after discharge, should be included.

Secondly, we will generate lay language by taking advantage of resources such as the Consumer Health Vocabulary (Doing-Harris and Zeng-Treitler, 2011; CHV, 2013), which maps medical terms to plain-language expressions. Additionally, we will pursue more sophisticated ex-

traction and rendering of rhetorical relationships among events and their outcomes (Mancini et al., 2007).

Last but not least, we will perform user studies, both controlled evaluation of our summaries while still at the development stage, and eventually longer-term assessments of whether our summaries engender better adherence to medications and better keeping of follow-up appointments, and ultimately, better health.

Acknowledgements

We thank three anonymous reviewers for their helpful comments. For partial financial support, we acknowledge award NPRP 5-939-1-155 from the Qatar National Research Fund, the UIC Department of Biomedical and Health Information Sciences, and the UIC Institute for Translational Health Informatics.

References

- Carol H. Cain, Estee Neuwirth, Jim Bellows, Christi Zuber, and Jennifer Green. 2012. Patient experiences of transitioning from hospital to home: an ethnographic quality improvement project. *Journal of Hospital Medicine*, 7(5):382–387.
- CDC. 2011. Hospital utilization (in non-federal short-stay hospitals). Centers for Disease Control and Prevention, <http://www.cdc.gov/nchs/fastats/hospital.htm>. Last accessed on 11/26/2013.
- 2013. Consumer Health Vocabulary Initiative.

- <http://www.layhealthinformatics.org/>. Last accessed on 5/19/2014.
- Barbara Di Eugenio and Nancy L. Green. 2010. Emerging applications of natural language generation in information visualization, education, and health-care. In Nitin Indurkha and Fred J. Damerau, editors, *Handbook of Natural Language Processing, Second Edition*, chapter 23, pages 557–575. CRC Press, Taylor and Francis Group, Boca Raton, FL. ISBN 978-1420085921.
- Kristina M. Doing-Harris and Qing Zeng-Treitler. 2011. Computer-assisted update of a consumer health vocabulary through mining of social network data. *Journal of Medical Internet Research*, 13(2).
- C. Friedman, L. Shagina, Y. Lussier, and G. Hripsak. 2004. Automated encoding of clinical documents based on natural language processing. *Journal of the American Medical Informatics Association*, 11(5):392.
- Albert Gatt and Ehud Reiter. 2009. SimpleNLG: A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation*, pages 90–93. Association for Computational Linguistics.
- N. Green, R. Dwight, K. Navoraphan, and B. Stadler. 2011. Natural language generation of biomedical argumentation for lay audiences. *Argument and Computation*, 2(1):23–50.
- C. Hallett. 2008. Multi-modal presentation of medical histories. In *IUI '08: Proceedings of the 13th International Conference on Intelligent User Interfaces*, pages 80–89, New York, NY, USA. ACM.
- G.M. Keenan, J.R. Stocker, A.T. Geo-Thomas, N.R. Soparkar, V.H. Barkauskas, and J.A.N.L. Lee. 2002. The HANDS Project: Studying and Refining the Automated Collection of a Cross-setting Clinical Data set. *CIN: Computers, Informatics, Nursing*, 20(3):89–100.
- Saad Mahamood and Ehud Reiter. 2011. Generating affective natural language for parents of neonatal infants. In *Proceedings of the 13th European Workshop on Natural Language Generation*, pages 12–21, Nancy, France, September. Association for Computational Linguistics.
- Clara Mancini, Christian Pietsch, and Donia Scott. 2007. Visualising discourse structure in interactive documents. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 89–92, Saarbrücken, Germany, June.
- NLM. 2009. UMLS Reference Manual. Technical report, National Library of Medicine, September. <http://www.ncbi.nlm.nih.gov/books/NBK9676/> (Last accessed on 12/09/2013).
2014. NNN: Knowledge-based terminologies defining nursing. <http://www.nanda.org/nanda-i-noc.html>. (Last accessed on 5/19/2014).
- Douglas P. Olson and Donna M. Windish. 2010. Communication discrepancies between physicians and hospitalized patients. *Archives of Internal Medicine*, 170(15):1302–1307.
- François Portet, Ehud Reiter, Jim Hunter, Somayajulu Sripada, Yvonne Freer, and Cynthia Sykes. 2009. Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence*, 173:789–816, May.
- Ehud Reiter, Roma Robertson, and Liesl Osman. 2003. Lessons from a failure: Generating tailored smoking cessation letters. *Artificial Intelligence*, 144:41–58.
- Anne Schneider, Alasdair Mort, Chris Mellish, Ehud Reiter, Phil Wilson, and Pierre-Luc Vaudry. 2013. MIME - NLG Support for Complex and Unstable Pre-hospital Emergencies. In *Proceedings of the 14th European Workshop on Natural Language Generation*, pages 198–199, Sofia, Bulgaria, August. Association for Computational Linguistics.
- Donia Scott, Catalina Hallett, and Rachel Fettiplace. 2013. Data-to-text summarisation of patient records: Using computer-generated summaries to access patient histories. *Patient Education and Counseling*, 92(2):153–159.
- Sandra Williams, Paul Piwek, and Richard Power. 2007. Generating monologue and dialogue to present personalised medical information to patients. In *Proceedings of the Eleventh European Workshop on Natural Language Generation*, pages 167–170, Saarbrücken, Germany, June.