SafeDischarge: A DSL-First Copilot for Safe, Patient-Specific Discharge Instructions

by

Steven Reyes

S.B. Computer Science and Engineering, MIT, 2025

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2026

Authored by:	Steven Reyes Department of Electrical Engineering and Computer Science May 15, 2026
Certified by:	Wojciech Matusik Professor of Electrical Engineering and Computer Science, Thesis Supervisor
Certified by:	Peter Szolovits Professor of Electrical Engineering and Computer Science, Thesis Supervisor
Accepted by:	Katrina LaCurts Chair, Master of Engineering Thesis Committee

SafeDischarge: A DSL-First Copilot for Safe, Patient-Specific Discharge Instructions

by

Steven Reyes

Submitted to the Department of Electrical Engineering and Computer Science
on May 15, 2026 in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

ABSTRACT

Writing discharge instructions is an important but time-consuming task for hospital clinicians, possibly leading to higher error rates and clinician burnout. Large language models (LLMs) offer the potential to speed up this workflow by synthesizing patient's electronic health records into discharge instructions, even personalizing them to a patient's comprehension level or language. However, existing methods lack clinical safety guarantees and still require extensive review, negating their benefits. This thesis explores a semi-automated data-driven approach to developing DischargeDSL, a domain-specific language (DSL) for discharge instructions in a clinician-verifiable format, and SafeDischarge, an LLM copilot to generate patient-specific DSL programs for discharge instructions that are more personalized and have less hallucinations. Our results show high recall in different discharge instruction topics, and demonstrate deterministic compilation to standardized discharge instructions, patient data citation ability, and automated dataset curation for prompt optimization. Further work is needed for discharge summary grammar development.

Thesis supervisor: Wojciech Matusik
Title: Professor of Electrical Engineering and Computer Science

Thesis supervisor: Peter Szolovits
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

I would like to thank my PIs, Prof. Wojciech Matusik and Prof. Peter Szolovits, for taking a chance on me with a research assistantship, as well as their continuous and unwavering support throughout the ups and downs of this project. This research study has given me a lot of valuable work experience with large language models and dataset curation that would not have been possible without their guidance.

I would also like to thank the EECS Undergraduate Office, particularly Katrina LaCurts, Ellen Reid, and Jessica Zdon-Smith, for being receptive of issues of the student body through the Undergraduate Student Advisory Group in EECS, and by extension, me as well. They have also been very helpful in keeping track with graduation progress and in supporting me in the MIT journey in general.

Finally, I would like to thank my friends, from both the Philippines and MIT, for making my experience in the last 5 years fun and memorable, and my family, for whom I would not be here without and for believing in me through to the end.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Clinical documentation takes up a large amount of clinician working hours [1], accounting for around half of an 11.4-hour workday [2]. Clinical documentation is complex, requiring clinicians to determine the past hospital course from previous clinicians' notes and to document pertinent medical findings, current hospital course, treatment plans, and discharge instructions in the patient's discharge summary. Consequently, meticulous documentation becomes time-consuming, so that 37% of patient's discharge instructions are delayed by at least a day [3], while also increasing opportunity for error and risk of burnout from administrative burden [4].

Large language models (LLMs) have the potential to reduce documentation burden, and thus improve patient care and clinician quality of life. LLMs, using their inherent clinical knowledge [5], can perform a large amount of clinical text summarization, such as compiling examination reports and synthesizing a coherent patient treatment course from multiple specialists, and have been shown to outperform clinicians in [6]. As such, when given the patient's health record as input, LLMs can generate patient-specific discharge notes and instructions, to be used as a first draft. Additionally, LLMs allow for more automated personalization, as they can tailor fit discharge instructions to the patient's background, e.g.,

comprehension level and native language, thereby improving instruction clarity.

Currently, the technology with LLMs is still far from the ideal of fully automating discharge summary generation. The limitations of the technology still require safety checks done by clinicians to vet for safety. However, care must be given to not overload clinicians with safety checks, otherwise, they are at risk of overlooking any potentially harmful errors that are more difficult to glean, especially when burnt out. In such cases, the adoption of the technology might increase the error rate, much more than if it hadn't been used at all.

1.2 Thesis Contributions, Summary, and Outline

This thesis is an attempt to bridge this LLM performance gap with the following contributions.

1. Exploring the development and evaluation of SafeDischarge, a grammar-based LLM copilot, as a potential avenue for safer and more patient-specific discharge instructions for coronary artery bypass graft (CABG) procedures.

2. Developing a semi-automated methodology to develop the grammar in a data-driven manner with clinician validation.

SafeDischarge constrains LLM outputs to the grammar of its custom domain-specific language (DSL), DischargeDSL. The DischargeDSL programs are composed of multiple fields to generate different parts of the output discharge instructions, which are filled in separately in SafeDischarge's modular LLM workflow, using only the necessary patient data to give an answer to each field. The entire pipeline outputs a DSL, whose contents are a concise, readable representation of patient-specific data that can be vetted by clinicians for transparency, and deterministically compiled to the final patient-facing discharge instructions through static discharge instruction templates with various reading levels and languages, to best suit the patient's background. This approach may preserve the LLM's benefits of saving time and reducing error, while also creating safe, transparent, and uniform discharge instructions.

This thesis also describes a semi-automated methodology to develop the DischargeDSL grammar, which uses a locally deployed LLM to examine the hospital data in the MIMIC-III and MIMIC-IV datasets. The main contribution of this methodology is an LLM pipeline that can validate actual patient data one-by-one with the grammar, so that the rarer cases not covered by the grammar can be surfaced. This allows for a more comprehensive method to compile the rarer patient cases, representing a major portion of the difficulty of writing discharge summaries, according to a clinician correspondent. With the rarer cases in hand, a clinician can be asked to review these rarer cases, and decide which ones are best to be implemented for automation to speed up time rather than obstruct clinicians with unnecessary review. This method has the advantage of being data-driven, rather than surveying clinicians for these rarer cases, which still risks being incomplete.

While most studies focus on discharge summary creation as the goal, this thesis mostly explores discharge instruction creation, a subset of discharge summaries. Discharge summaries have some parts that are less structured than discharge instructions, and the list of rare cases in discharge instructions to be presented to clinicians for review is more likely to be doable than discharge summaries. Hence, discharge instructions were given more focus in this thesis, with an extended exploration into discharge summaries afterwards.

The remainder of this thesis is structured as follows. **Chapter 2** describes related literature motivating this thesis. **Chapter 3** describes the methodology to develop the DischargeDSL grammar and the SafeDischarge LLM copilot. **Chapter 4** describes results obtained from the process, and concludes with future directions.

Chapter 2

Related Work

2.1 LLM-Generated Discharge Summaries

LLM-generated discharge summaries tend to have 1.5-1.6 times as many errors per summary compared to clinician-written ones [4, 7], but errors usually have low overall harmfulness scores [7]. When categorized by type, errors of omission happen more frequently than errors of hallucination [4, 8]. Majority of errors have low potential for harm [7, 9], especially errors of omission [8]. No significant difference in LLM- and physician-generated narratives is found when rated by blinded reviewers on overall quality and preference [7].

While hallucinations occurred less often, they are about three times more likely to be classified as major compared to omissions [8], where a major error is defined as an error that can affect the management and diagnosis of the patient more than when otherwise correct. LLM- and physician-generated narratives are also shown to contain a similar number of hallucination errors [7]. Different parts of the discharge summary have varying levels of hallucination and omission rates. Hallucinations most often occur in the future-facing plan section [8, 9], while omissions most often occur in past- and present-facing sections, such as current issues [8], physical examination results, and history of presenting complaint [9].

While LLM-generated discharge summaries tend to be more concise and coherent, they

are usually less comprehensive than clinician-written summaries [7]. Specifically, only 60% of medical professionals rated LLM-generated summaries as sufficiently comprehensive [10].

Some of these issues are addressed by the DischargeDSL compiler's validation checks. As will be explained in more detail in **Chapter 3**, hallucination errors are addressed by prompting the LLM to list the verbatim evidence it uses for its answer, and then checking it against the input, reprompting the LLM upon encountering hallucinated evidence. This same procedure can be helpful in reducing errors of omission too, but a comprehensive method for minimizing omissions is still lacking.

2.2 Advanced Prompting and Agentic Approaches

While most studies evaluate LLM performance via naive prompting methods, where all the free-text notes found for a single patient's stay are fed as input to the LLM, with some prompt engineering and few-shot examples, a few studies have explored more in-depth with more advanced methods of prompting and with LLM agents.

When trying different prompting techniques, such as structured prompting, function calling, atomization of free-text into lists of facts, JSON-based outputs, and templating, the highest performance improvements were found when instructing LLMs to produce a JSON output, include negative findings beside positive ones, and allow an unknown status category as a possible answer for problems not mentioned in transcripts, achieving only 2 hallucinations among major errors in 25 notes [8]. Another study converted hospital discharge summary writing guidelines to a JSON schema, and fed into the LLM alongside patient data, which resulted in around 2.5 major omissions and 1 hallucination per summary for 53 summaries [11].

Among agentic approaches, one study used two agents: one for continuous refinement of discharge summaries, where the current discharge summary draft is refined with any new notes written during the day, over the course of the hospital stay, and another one to

add citations to claims in the discharge summary, removing any sentences without a valid citation [12]. This reduced hallucination rates to 2%, while omissions rates were measured to be 25%. The study also found that 88% of errors were rated not harmful by physicians, while physician burnout scores decreased. A different study used four different agents, one to summarize patient data, one to draft a discharge summary, one to detect hallucinations, and one to revise drafts and remove hallucinations [13]. This pipeline resulted in better completeness, coherence, and factuality than the baseline naive prompting of LLMs, though evaluations are done via LLM-as-judge, and no comparisons were made against the ground truth of clinician-written summaries. However, no other study has been found doing a similar agentic approach based on a data-driven grammar as in SafeDischarge.

2.3 MIMIC-III and MIMIC-IV Databases

The DischargeDSL grammar for discharge instructions is developed using the MIMIC-IV electronic health record dataset [14], hosted on PhysioNet [15]. The MIMIC-IV dataset includes data from around 360,000 patients' electronic health records at the Beth Israel Deaconess Medical Center in Boston, including structured data such as patient information, diagnoses, treatments, procedures, medications, lab measurements, and ICU chart data, as well as free-text clinical notes like radiology reports and discharge summaries in the MIMIC-IV-Note extension, that have all been de-identified for privacy [16]. The MIMIC-IV dataset is built upon the previous MIMIC-III dataset [17, 18]. The MIMIC-III dataset has a more comprehensive list of free-text notes per patient than MIMIC-IV's radiology reports and discharge summaries. As such, it is used primarily for the development of the grammar for discharge summaries.

2.4 Clinical Domain-Specific Languages

There are existing DSLs used for clinical settings, such as the Clinical Quality Language (CQL) Specification [19] and the Arden Syntax [20]. However, these DSLs in particular are primarily used for clinical decision support (CDS) systems that assist doctors, rather than create patient-facing discharge instructions. Other DSLs that support clinical documents exist, such as the Fast Healthcare Interoperability Resources (FHIR) [21] and the Clinical Document Architecture (CDA) [22]. While these DSLs can express discharge summaries through their grammars [23], they are intended for broad scope, meant for use by "patients, healthcare providers, and other stakeholders" [24]. They include numerous metadata fields that make it less readable for clinicians and LLMs. DischargeDSL alleviates this by only including necessary fields that doctors need to be as concise as possible, reducing documentation burden.

2.5 Domain-Specific Languages for LLMs

DSLs have been shown to improve LLM-generated output accuracy in other domains. Graphics-based DSLs allows for more accurate and realistic outputs, such as in MoVer that improves accuracy of LLM-generated animations from 58.8% to 93.6% by enforcing motion validation [25], and in VLMaterial that uses a Python-based DSL to generate procedural materials in order to design photorealistic material appearances, preserving perceptual similarity between input images and output materials [26]. In the domain of physical modeling, MetaDSL has a grammar that encodes diverse metamaterial designs that is human-readable and machine-parseable [27], while Structured Parametric CAD Code (SPCC) is a DSL utilized by CAD-Llama, which is a framework for generating parametric 3D CAD models [28]. This shows a promising path for DSL-based LLM pipelines for use in the hospital.

Chapter 3

Methodology

3.1 Discharge Instructions

3.1.1 Data Pre-processing

The MIMIC-IV dataset is first filtered for patient cases who had undergone CABG surgery and have discharge instructions present. Each patient is assigned a unique `subject_id`, with different hospitalizations assigned a unique `hadm_id`. To filter for CABG cases, the `procedures_icd` table of MIMIC-IV is used to filter for ICD procedure codes and versions that describe a CABG procedure. The ICD lists multiple codes for CABG variations, all under the same prefix for CABG, as listed in Table 3.1. This gives a tabulation of `(subject_id, hadm_id)` pairs. From here, the `discharge.csv` table of MIMIC-IV-Note is used to filter for pairs that have any discharge note present. This process gives 7245 CABG patient admissions, and 5629 of them with discharge notes.

The bulk of the workload is prompting an LLM patient-by-patient, so each table of MIMIC-IV and MIMIC-IV-Note is split by `subject_id` into subtables as `.csv` files in a patient's own subdirectory, except for non-patient-specific tables such as item ID definitions.

For discharge instruction generation, the target output is the discharge instruction section of the discharge summary, and the rest of the patient data, including the tabular data in

ICD code prefix	ICD version	Text description
361	9	(Aorto)coronary bypass for heart revascularization
0210	10	Bypass Coronary Artery, One Artery
0211	10	Bypass Coronary Artery, Two Arteries
0212	10	Bypass Coronary Artery, Three Arteries
0213	10	Bypass Coronary Artery, Four or More Arteries

Table 3.1: ICD codes and versions for CABG procedures

MIMIC-IV, radiology reports, and the rest of the discharge summary, serve as the input. The discharge instructions are extracted via regex, matching for text between the "Discharge Instructions:" header and the succeeding section header.

3.1.2 Model Prompting Setup

Since MIMIC-IV contains real-world patient data that cannot be shared with third parties even though it is de-identified, a locally deployed LLM setup is used for prompting. The open-source model used is gpt-oss-120b [29], deployed via llama.cpp [30] as a local server, with 0.1 temperature to minimize hallucinations, medium reasoning effort, and the maximum context window size of 131072.

Prompting is initially done via the OpenAI Python package, prompting the local server similar to using the OpenAI API. To optimize the prompts given, the DSPy Python package is used to declaratively specify model inputs and desired outputs, and to optimize prompts using its built-in optimizers like GEPA reflective prompt evolution [31, 32].

In the latter half of this research work, MIT released Parley, a secure MIT-hosted platform running the latest OpenAI, Anthropic, and other LLMs while still preserving data privacy [33]. Parley is used for one-off prompting about patient data, such as prototype grammar generation or clarifying questions about patient data. Patient-by-patient prompting is not possible due to a lack of an API.

3.1.3 Prototype Output Grammar Development

DischargeDSL is developed by first creating a standardized grammar for the outputs, i.e., the discharge instructions, from actual MIMIC-IV discharge instructions, before developing the DischargeDSL grammar itself. This is to capture all possible output variations and to standardize the outputs. The data-driven grammar development process is done in three stages: creating a prototype grammar from a corpus of MIMIC-IV discharge instructions, standardizing MIMIC-IV discharge instructions one-by-one against the grammar, and noting outlier sentences not captured by the grammar.

To create the prototype grammar, first, the local LLM is prompted to output a list of sentences most commonly found in the discharge instructions, while also passing as many of the regex-extracted discharge instructions that can fit in one prompt. This resulted in 10 promptings of the LLM, with around 600 discharge instructions each. The 10 output lists of sentences are then handpicked into a single list of sentence topics, the main ones being:

1. **Shower and incision care**: Please shower daily including washing incisions gently with mild soap, no baths or swimming, and look at your incisions. Pat incisions dry, do not rub.
2. **Topical product restriction**: No lotions, creams, powders, or ointments to incisions.
3. **Daily weight and temperature log**: Each morning you should weigh yourself and then in the evening take your temperature, these should be written down on the chart.
4. **Driving restriction**: No driving for one month or while taking narcotics.
5. **Lifting restriction**: No lifting more than 10 pounds for 10 weeks.
6. **Office contact**: Please call cardiac surgery office with any questions or concerns.
7. **Female-specific incision care**: Females: Please wear bra to reduce pulling on incision, avoid rubbing on lower edge.
8. **Infection monitoring**: Monitor wounds for redness, drainage, or increased pain.
9. **Fever**: Report any fever greater than 100.5.
10. **Weight gain**: Report any weight gain of 2 pounds in 24 hours or 5 pounds in 1 week.

11. **Shoulder movement**: Encourage full shoulder range of motion, unless otherwise specified.
12. **Prevena dressing instructions**: The Prevena wound dressing should be left on for a total of 7 days post-operatively...
13. **Medications**: Take all medications as prescribed.
14. **Fluid restriction**: Fluid restriction: 1.5 L per day until seen.
15. **Sunscreen**: Use sunscreen on incision if exposed to sun.
16. **Coumadin/Warfarin**: You are taking coumadin (a blood thinner) for atrial fibrillation, goal INR of 2.0-2.5. Coumadin dosing will be managed by Dr. ____.
17. **Binders**: Wear sternal/abdominal/chest binder at all times except for showering.
18. **Other medication-specific**: Medications vary, such as cipro antibiotics, lasix, potassium, amiodarone, insulin, metformin, etc.
19. **Glucose monitoring**: Monitor finger stick blood sugars twice daily.
20. **Smoking restriction**: Please stop smoking. Information was given to you on admission regarding smoking cessation.
21. **Alcohol restriction**: Do not drink alcohol while taking narcotic pain medicine.
22. **NSAID restriction**: Please DO NOT take any non-steroidal anti-inflammatory medications (NSAIDs such as celebrex, ibuprofen, advil, aleve, motrin, etc).

Each sentence topic is inspected and confirmed by searching in the corpus for the verbatim keywords specified by the LLM. For example, sentence topic 16 has many wording choices, and many were found via searching for "coumadin" or "warfarin", ignoring capitalization. Sentence topics that were redundant or not found in the corpus were removed. Moreover, for sentences with vast variations, those are noted as well in the finalized sentence list, with variations marked as "Sometimes", "Rarely", and "Only once". Sentence topics 1 to 15 are mostly the same verbatim text repeated between discharge instructions, while 16 to 22 have more variation. Among the quickly inspected discharge instructions for sanity checking, sentences 1 to 10 are present in nearly every one of them, 11, 12, 13 and 15 occur sometimes, and some items like 17 and 20 to 22 appear only around 1 to 10 times.

Then, the prototype output grammar is created from this sentence topic list by prompting the LLM to create the grammar in the Extended Backus–Naur form (EBNF) specification, using the topic list and the same corpus of discharge instructions, as before. A previous grammar-generation attempt was performed with the corpus alone, but that grammar only included sentence items 1 to 10, with the rest of the items only designated as a catch-all `otherInstructionLine` grammar symbol.

3.1.4 One-by-one Grammar Validation

The prototype grammar is then validated against each patient's actual discharge instructions. This is to check for any examples that are not represented in the grammar, from which either the grammar can be updated, or the sentences not represented in the grammar can be added to a list of rare sentences. Figure 3.1 shows a flow chart for this validation pipeline. For each MIMIC-IV patient's discharge instructions, the local LLM is prompted with the EBNF grammar, the original discharge instructions, and a prompt instructing the LLM to derive the standardized discharge instructions from the given text using the grammar.

The LLM-generated standardized discharge instructions are then passed to two parallel LLM checks: grammar validation and semantic preservation. The former check makes sure the generated standardized discharge instructions can be parsed correctly as a DSL program of the given grammar. The latter check asks the LLM if there is any added or removed semantic content in the standardized discharge instructions from the original. Both checks are done in parallel, with each prompt only needing the conversation history of the previous DSL program generation.

Here, the grammar is slightly modified, where the catch-all `otherInstructionLine` grammar symbol is removed. This is to ensure the standardized discharge instructions successfully output by the LLM can only come from the prescribed discharge instruction lines of known format. Any unknown sentences are caught as LLM generation errors, rather than transparently passed to the standardized output, without prior human standardization.

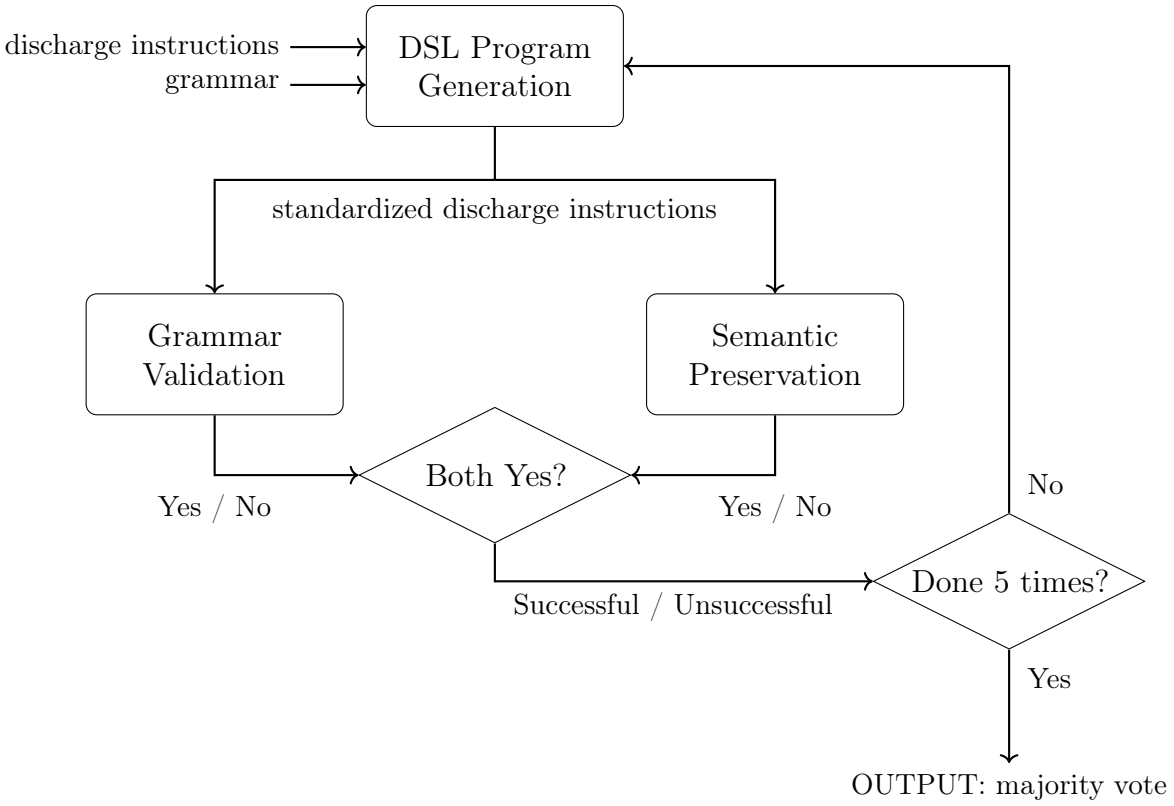

Figure 3.1: Pipeline for generating and validating standardized discharge instruction with LLMs. Boxes represent LLM calls, while diamonds represent deterministic computer logic.

Furthermore, the first prompt of DSL program generation also includes a note not to be too strict with preserving semantic content. Otherwise, the LLM would consider the following pairs as different: "for one month" vs. "for approximately one month", and "Report any weight gain and fever" vs. "Report any weight gain or fever". An example of such pair is also included in the prompt.

The first few examples are prompted by hand, since the ones at the beginning are most likely to show what minor edits the grammar needs. Afterwards, the pipeline is done automatically for the remaining patients. To increase accuracy, the pipeline is repeated 5 times per patient, and the majority outcome is taken as the final result for this patient. All LLM prompt outputs were also logged to disk while iterating over the patients for persistence.

Outlier category	Keywords
Prevena	prevena
Sunscreen	sunscreen, sunblock, suntan, sun cream
Washing, not showering	shower (negative match)
Binder	binder
Shoulder movement	shoulder, range of motion
Fluid restriction	fluid
No alcohol	alcohol, drink
No smoking	smoke, smoking, tobacco, nicotine
NSAIDs	nsaid, ibuprofen, advil, motrin, naproxen, aleve, aspirin, celecoxib, celebrex, diclofenac, voltaren, meloxicam, mobic
Coumadin	coumadin, warfarin, inr
Insulin	insulin, metformin, lantus, sugar, glucose
Diuretics	diuretic, diuresis, lasix, furosemide
Electrolytes	electrolyte, potassium, magnesium, k+, mg
Blood pressure	blood pressure, hypertension, losartan, lisinopril, beta blocker, betablocker, ace

Table 3.2: Keywords used for outlier categories

3.1.5 Noting Outlier Sentences

The two previous stages explain the automated portion of the development of the discharge instruction grammar. Now, this stage represents the manual curation of outliers from the unsuccessful LLM generations. Since the number of unsuccessful generations were in the order of a hundred, these were inspected until the outliers were mostly repeating, with no new types of outliers appearing. In practice, the number of cases manually checked was 265.

Each unsuccessful generation is inspected, deciding whether the grammar needs a small edit to accommodate the original discharge instructions, or that it's an outlier to be logged. Small edits to the grammar are usually minor wording differences, while outliers are categorized by topic. Then, after processing through the unsuccessful generations, a rough estimate of the total number of cases each outlier category appears in is obtained by searching for relevant keywords across the entire corpus, as shown in Table 3.2. The same is done for sentence categories already in the discharge instruction grammar, but only sometimes occur. Each discharge instruction is converted to lower-case, to prevent mismatches due to capitalization.

This then gives a list of sentence topics that are not constant in every discharge instruction, alongside their frequency rates in the original corpus. A threshold frequency rate can be set, possibly through prior clinician consultation, below which the LLM does not offer any speed-up due to editing overhead. From these rare topics, as well as the occasional topics already in the prototype output grammar, the DischargeDSL grammar itself can be developed. This requires knowing what variations occur in the corpus, and what information is needed from each patient to decide which specific variation to use, which is provided by this collection of occasional and rare sentence topics.

3.2 DischargeDSL and Outputter

After the output grammar is finalized, the program to translate the intermediate DischargeDSL program to the standardized discharge instruction output can be implemented. The output grammar accommodates multiple possible word choices for the same ideas, so these wordings are first stripped down and standardized to only a handful of meaningfully different ones. For example, different instructions have different lists of red flags to look for to know when to call the clinician, such as nausea, yellowing of skin or eyes, or feeding tube becoming clogged, so the outputter program standardizes this list and outputs all red flags found.

The outputter program thus takes in three different types of inputs: custom settings that the clinician can set, templated text fields, and those based on patient-specific data. First, the custom settings allow for different wordings that appear in the corpus, while still enforcing some standardization. Most of the text in the corpus follows a standard wording, but there were a few that were simpler than usual, and a few that were more detailed. Thus, for each sentence topic, at most three wording choices are chosen to accommodate these different comprehension levels. Furthermore, some sentence topics had different choices of clearing conditions. For example, some were told not to drive until cleared by the surgeon, while others were restricted to a fixed period of one month. The clinician can set which of

these three comprehension levels and which clearing condition to use for each sentence topic, depending on what they feel is appropriate for the patient.

The word choices for each parameter setting are stored in a YAML file, for readable and editable templates for clinicians without needing to change the underlying code. Currently, the YAML file is in English, and was translated into other common languages in US hospitals, such as Spanish and Mandarin Chinese, via ChatGPT-5.4, but these translated versions still necessitate verification from clinicians who speak these languages before they can be used in actual hospitals. However, this does show how the implementation can easily scale to other languages as needed.

Second, templated text fields involve data that isn't patient-specific, but still is usually constant for the clinician or hospital. These include phone numbers and other contact information, follow-up appointment physicians' names and offices, follow-up schedules, emergency hotline numbers, etc.

The remaining input fields then comprise the patient-specific data that is embodied in a DischargeDSL program to determine the variation of discharge instructions to use. Most sentence variations are whether or not the sentence is included in the final discharge instruction, so it can be represented in DischargeDSL as a boolean value. For sentence variations, a few examples are fed into an LLM like ChatGPT-5.4, asking it what patient data is required to determine which variation to use. Then, those data fields are also placed into the DischargeDSL grammar.

However, some meaningfully different variations of sentences could not be narrowed down to the most appropriate one with patient data, such as the maximum lifting restriction (5 vs. 10 pounds), how long to restrict lifting for (6 weeks vs. 10 weeks vs. 2 months), or how long to restrict driving for (4 weeks vs. 6 weeks). When prompting ChatGPT-5.4 for how these variations are decided, it's most likely a clinician preference rather than encoded in patient data. As such, these variations are standardized to a single value in the outputter, though a clinician may choose to edit these values in the YAML template file.

3.3 DischargeDSL Program Extraction from Patient Data

Topic fields	Input patient data
Prevena included	`admissions`, `patients`, `diagnoses_icd`, `procedures_icd`, `chartevents`
Shoulder movement encouraged	`admissions`, `patients`, `diagnoses_icd`, `procedures_icd`, `radiology`
Coumadin included, indication, INR	`admissions`, `patients`, `diagnoses_icd`, `procedures_icd`, `labevents`, `pharmacy`, `prescriptions`
Diuretics included	`admissions`, `patients`, `diagnoses_icd`, `procedures_icd`, `labevents`, `pharmacy`, `prescriptions`

Table 3.3: Input patient data tables used for each topic's fields

This section describes how an LLM is used to determine the correct DischargeDSL fields for each sentence topic from the other MIMIC-IV patient data. The fields of each topic are divided into their own workflows, to narrow down the input patient data to only what's necessary for those fields, as elaborated in Table 3.3. The input patient data tables are decided by prompting ChatGPT-5.4 for the relevant inputs for each sentence topic.

Figure 3.2 shows how the patient data is processed into the output DischargeDSL fields. First, the structured patient data tables are processed into a string LLMs can ingest. Each row of the input tables is transformed into comma-separated fields in one line, each taking the form `column_name: value`, so the LLM does not need to remember the ordering of the columns when processing through the data. These rows are annotated with a tag and row index, e.g., rows in `labevents` are annotated with `[LAB:n]`, `diagnoses_icd` with `[DX:n]`, etc. Tables such as `admissions` and `patients` have only one row for the patient stay, so each field is given its own line. Meanwhile, `labevents`, `chartevents`, `procedures_icd`, and `diagnoses_icd` refer to item IDs defined in `d_labitems`, `d_items`, `d_icd_procedures`, and `d_icd_diagnoses`, respectively, so these tables are joined by item ID. To save on tokens, datetimes in the two largest tables, `labevents` and `chartevents`, are expressed in MM-DD HH:MM format, and only the relevant lab items, ICU chart items, and medications were selected for the input, as listed in Table 3.4. These filter items are chosen by prompting

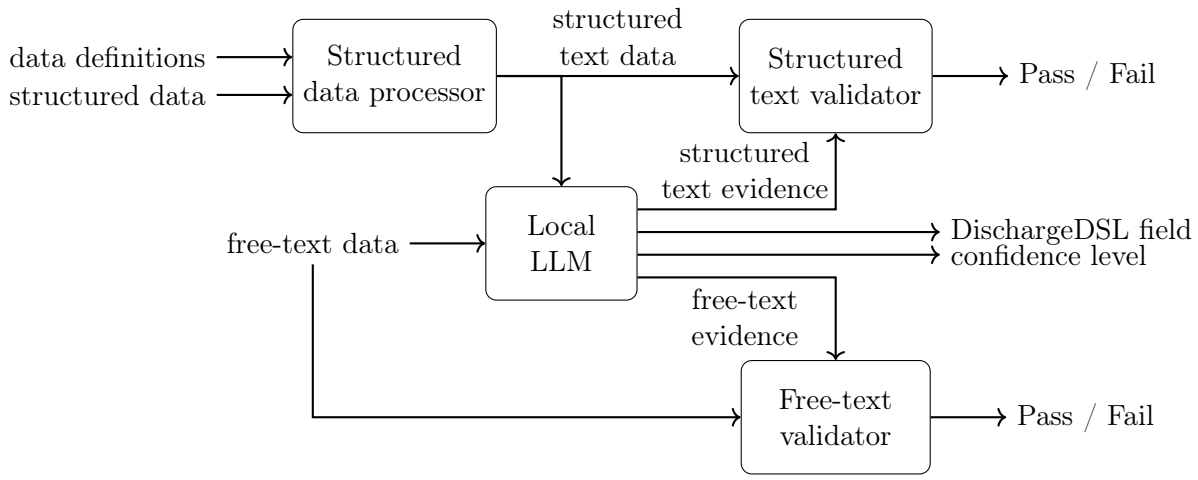

Figure 3.2: Workflow to generate and validate DischargeDSL programs extracted from MIMIC-IV patient data

ChatGPT-5.4 for a comprehensive list of item IDs for the given patient data table that can answer the topic's DischargeDSL fields.

Next, the structured text string, together with the MIMIC-IV free-text data such as radiology reports, are passed to the LLM. Using DSPy, the LLM is asked to give the following outputs: the desired DischargeDSL field, a confidence level among low, medium, and high, and both structured and free-text evidence it uses to support its answer, each with their own reasoning field for a chain-of-thought style of prompting. It cites the free-text evidence by giving the text verbatim, while it cites the structured text by giving the table's tag, row index, and verbatim text. These are so that the validators can verify that the LLM cited its reasoning correctly, in order to reduce hallucination errors.

The validators use the RapidFuzz Python package to fuzzily match the LLM-cited text and the original reference text. For structured text, it uses `token_set_ratio` with a threshold of 90%, since it's acceptable if the LLM changes the order of the fields, as long as they have the same content, modulo rearrangement. For free-text, `partial_ratio_alignment` is used with a threshold of 90%, which gives a text span in the reference text that matches the citation. If all citations are verified to be part of the text, then the workflow gives an output, with the text evidence it cites, to the clinician for review. The structured table tags and row indices,

Topic	Table	Selected items
Prevena	`chartevents`	Incision closure, appearance, drainage, drainage amount, dressing, cleansing packing, Dressing status, Skin treatment, site, type, cleanse, drainage, drainage amount, wound base, integrity, condition
Coumadin	`labevents`	INR(PT)
Coumadin	`pharmacy`	Coumadin, Warfarin
Diuretics	`labevents`	Potassium, Magnesium, Creatinine, Urea nitrogen, Sodium, Chloride, Bicarbonate
Diuretics	`pharmacy`	Furosemide, Lasix, Bumetanide, Torsemide, Metolazone, Chlorothiazide, Hydrochlorothiazide, Spironolactone, Eplerenone, Amiloride, Triamterene, Potassium chloride, KCl, Potassium phosphate, Potassium acetate, Magnesium sulfate, Magnesium oxide, Magnesium chloride, Magnesium gluconate, Lisinopril, Enalapril, Captopril, Ramipril, Benazepril, Losartan, Valsartan, Irbesartan, Candesartan, Sacubitril, Trimethoprim-sulfamethoxazole, Amiodarone, Sotalol, Dofetilide, Digoxin

Table 3.4: Selected input items for each topic and patient data table

and the free-text spans can be passed to a front-end site that highlights the evidence, for the clinician to validate the reasonability of the LLM. Otherwise, DSPy's Refine module asks the LLM to repeat the LLM generation. It has a reward function, whose feedback is given to a separate LLM call, so that it can provide additional insight to the next iteration as to why it failed.

The prompts for generating DischargeDSL fields are optimized by using DSPy's GEPA optimizer, which uses a dataset and a evaluation metric to choose the best prompt from a set of candidates. These candidates are generated by LLM as well, by asking it to reflect on how it could have better performed the task. For each topic, the dataset is created by prompting the unoptimized DSPy program with an LLM over randomly chosen patient data, until the DischargeDSL field boolean value that specifies whether to include the sentence or not, has at least 30 true and 30 false cases. Each case's actual true or false value is then obtained, to get a confusion matrix for each topic. Then, to get the final dataset, at most 10 data points are chosen from each cell of the confusion matrix. It's likely that the actually true cases (true

positives and false negatives) don't have 10 in each case, because of how rare true cases are in the corpus, so more data points are sampled from the unprompted actually true cases, until there are 20 of them in total. Then, more samples are taken from the naively prompted data points until 40 total data points are achieved. These sets of data are split equally into training, validation, and test sets for GEPA to use.

The metric used in GEPA checks whether the predicted boolean value to include the sentence or not, matches the actual value, giving a 1 for a match, and 0 otherwise. For text fields such as the indication for Coumadin prescription, an LLM-as-judge evaluation is used to check for semantic equivalence. The confidence level is not used for optimization, but is only used to determine whether to pass the LLM outputs to the clinician for review. Then, after optimization, a more comprehensive evaluation of the optimized program with 400 randomly sampled data points each is done, giving a final confusion matrix for each topic.

3.4 Discharge Summaries

Although the original intention of this project was to focus only on the discharge instruction part of discharge summaries, on studying these, we find that they are relatively highly structured, with the need for specific instructions dependent on only very specific parts of the overall patient record. Therefore, we decided to enlarge the scope of our work to try to create a similar analysis on the entire discharge summary, which contains many more components, including free-text sections such as the patient's presenting state, past medical history, family medical history, their hospital course, and discharge instructions, as well as more structured sections such as their medications on admission, during their hospital stay, and at discharge, and also tests and therapeutic procedures performed during hospitalization. However, with a larger scope of project and more free-text sections, results involving discharge summary generation are not as comprehensive as discharge instructions due to lack of time.

To gain access to free-text data used during hospital stays, which contain most of the

desired information for the discharge summary, the MIMIC-III dataset was used, in order to create the discharge summary grammar, where free-text notes are found in the `noteevents` table. The MIMIC-III notes are also pre-processed to be filtered for CABG patients with discharge summaries, and split into directories by patient `subject_id`.

To create the discharge summary grammar, the `noteevents` table is filtered for those with category of `"Discharge summary"`, and uploaded to MIT Parley's ChatGPT-5.4 to get a prototype grammar. The LLM is prompted to avoid expressing the grammar simply as a catch-all grammar symbol such as `note: WORD+`, but instead to list the common sentence formats first, while leaving the catch-all grammar symbol as a last resort. This is to capture as much structure in the free-text sections as possible. The grammar is then passed to Anthropic's LLM Claude to standardize 10 discharge summaries from different patients using the prototype grammar and output as JSON, in order to refine the grammar for any source material it didn't originally support. Usage of these models are compliant with PhysioNet's agreements on usage of MIMIC-III data [33, 34].

Afterwards, to create a DischargeDSL prototype, the rest of the patient data of the 10 patients are uploaded to Claude, while being prompted to inspect the 10 standardized discharge summary JSON files, and to split the files into two JSON files each, one containing the fields that can be deterministically filled in by the structured patient data, and one containing fields that have to be generated non-deterministically. The non-deterministic fields then comprise the bulk of DischargeDSL's grammar. To generate each field of DischargeDSL, the common sentence formats are done first. These common sentences usually contain template text, in between which some placeholder values are to be filled in. These placeholder values are filled in by prompting with a local LLM for these values. Specifically, for each common sentence format, ChatGPT-5.4 is prompted to know which MIMIC-III tables the correct placeholder value can be found in. Then, a local LLM is prompted with the selected input patient data to fill in the values of the DischargeDSL fields. Afterwards, to deal with any catch-all sentences missed, the common sentences are collected for the discharge summary

section, and prompted in the local LLM alongside the entire patient data, to refine the section with any other important details. Then, the discharge summary is finalized by collecting all the generated sections, and also adding in the deterministic fields, obtained by a static parser.

Since the benefit of discharge summary generation mainly lies in the unusual patient presentations and operations that are mostly ad hoc, the evaluation method is done to measure completeness of these unusual findings. Specifically, a list of topics is obtained for the predicted and actual discharge summaries via local LLM prompting. Then, they are compared semantically, again by local LLM, to find the matches. The topics that match are considered true positives, the topics omitted in the prediction but are present in the actual are considered false negatives, and the topics hallucinated in the predicted but absent in the actual are considered false positives.

Chapter 4

Results

4.1 Discharge Instruction Grammar

After pre-processing MIMIC-IV data, there are 5629 CABG patients with discharge instructions. The automated pipeline for discharge instruction grammar generation and validation was executed over 3648 patient discharge instructions. Among them, 51% had a successful pipeline generation and validation of standardized discharge instructions, 46% were unsuccessful, and 3% were skipped due to the patient expiring or lack of discharge instructions. Though not implemented for lack of time, the success rate could potentially be improved with DSPy. Its Refine module could have been used to get progressively better outputs over the 5 iterations, as it uses its own LLM call to refine the prompt instructions to the LLM for the task given.

Among the unsuccessful LLM generations, the first 265 of them were manually inspected. Among them, 25% were truly impossible to be represented by the grammar, because of an outlier sentence, while the remaining 75% were false negatives. Most of these false negatives were hallucinations, where during the comparison of semantic content between the standardized DSL program and original, the LLM changes part of the standardized DSL output, miscopying the text from its own generated answer, and thus thinking that the DSL

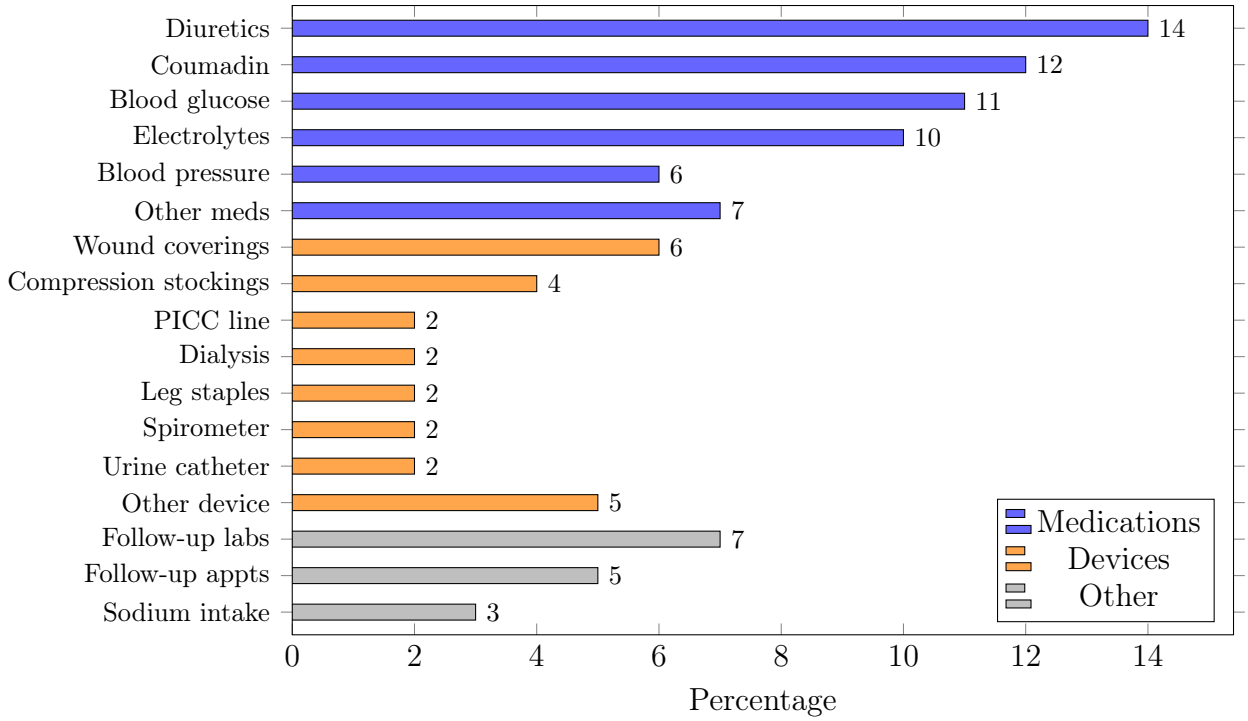

Figure 4.1: Percentage counts of outlier sentence topics in manually inspected outlier MIMIC-IV CABG discharge instructions, grouped by category

program has some information added or removed. Besides that, many of the false negatives were due to added or removed qualifiers, e.g. "approximately/at least one month" vs. "one month", or "Gently pat dry" vs. "Pat dry". This issue is supposedly addressed by the included note and examples in the prompt to not be strict with preserving semantic content, yet the LLM still makes these errors.

There were 100 outlier sentences among the 265 manually inspected discharge instructions from unsuccessful generations. Their types and percentage occurrences among these manual inspections are listed in Figure 4.1. These types include medication-specific ones such as diuretics/water pills, coumadin/blood clotting, insulin/blood glucose, electrolytes, blood pressure medications, and other, and also include device-specific ones such as wound dressing removal and replacement, leg compression stockings for edema, and other devices.

Those involving devices were not included in the DischargeDSL grammar, since there was no repetition between sentences, and no coherent formatting could be drawn from these

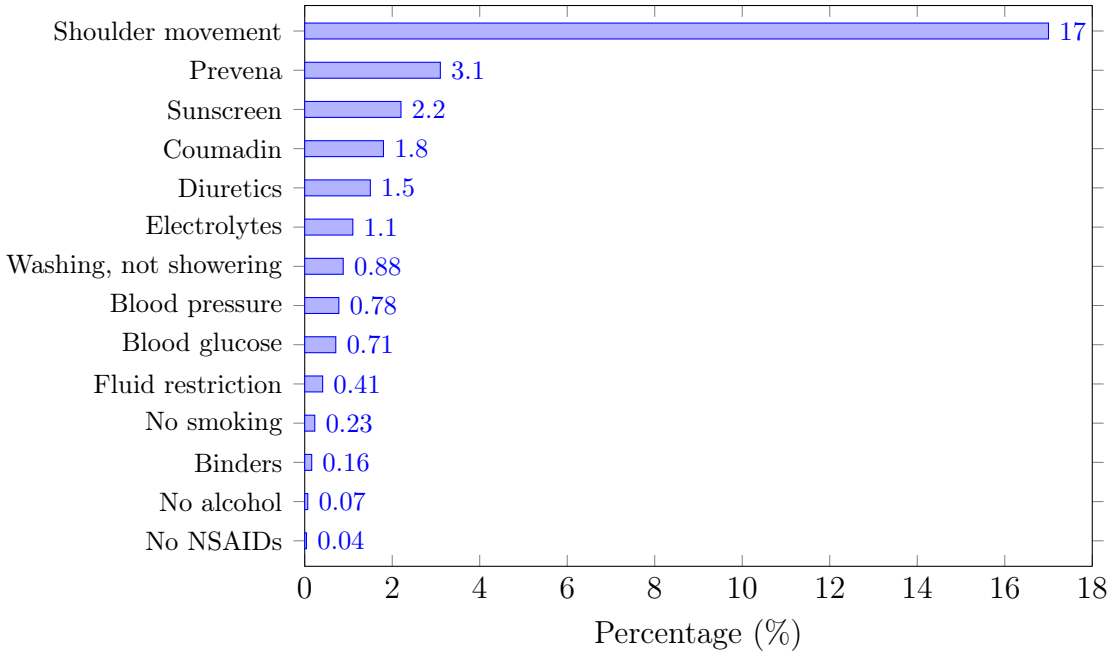

Figure 4.2: Percentage occurrences of occasional and outlier sentence topics in entire MIMIC-IV CABG discharge instruction corpus

sentences. In practice, these sentences are better off manually done by the clinician. The most common of these rare topics, as well as the occasional topics already in the grammar, such as Prevena dressing instructions, sunscreen, binders, are counted for their estimated total occurrences in the entire discharge instruction corpus in Figure 4.2. Among them, only the instruction encouraging shoulder movement has a substantial frequency rate, at 17%, the rest at 4% and lower. In fact, some sentences in the prototype output grammar occur less often than the outlier sentences found later on, such as washing instead of showering, fluid restriction, binders, and no NSAIDs. These sentences can also be removed from the final outputter due to this rarity.

The 5 most common topics here were chosen to be part of the DischargeDSL. For the sunscreen instruction in particular, since there is only one variation telling patients to use sunscreen on incision when exposed to sun, it was always included in the output, for simplicity. Prevena dressing and shoulder movement instructions also always had one single variation, so these two topics only need one boolean value each.

Topic	TP	FP	FN	TN	Recall	Precision	Acc
Prevena	15	142	0	231	1.00	0.10	0.63
Shoulder	50	81	0	207	1.00	0.38	0.76
Coumadin	6	32	2	260	0.75	0.16	0.89
Diuretics	5	356	0	30	1.00	0.01	0.09

Table 4.1: Recall, precision, and accuracy metrics for each topic in DischargeDSL's grammar

For coumadin and diuretics, the variations seem to mostly lie in the stopping condition for taking the medications. Some give a fixed dosage to take over a fixed time, some tell the patient to take a medication until a certain metric goal is reached (e.g. INR for coumadin and pre-operation weight for diuretics), and some say to follow-up with the clinician for further instructions. These variations seem to be mostly clinician preference, rather than based on patient data, so these cannot be encoded in the DischargeDSL grammar.

This leaves only a few fields for coumadin and diuretics. The latter only has a boolean field for inclusion/exclusion. From prompting ChatGPT-5.4, coumadin dosing depends on the indication that necessitates this medication, which also determines the desired INR target range. The grammar is thus:

Listing 4.1: Finalized DischargeDSL grammar

```
has_prevena: bool
encourage_shoulder_movement: bool
coumadin_instruction_needed: bool
coumadin_indication: str
coumadin_inr_goal: tuple[float, float]
diuretics_instruction_needed: bool
```

After creating a DSPy program, curating a dataset, and prompt-optimizing with GEPA for each sentence topic of DischargeDSL, Table 4.1 shows the evaluation metrics of the prompt-optimized DSPy programs for each topic across 400 randomly sampled patient data. For these metrics, recall is best to be optimized on, as it's worse if an instruction is omitted when it should be included, rather than the other way around, because the former allows a clinician to review whether to include this instruction or not.

Prevena and shoulder movement instructions have 100% recall, while coumadin and

diuretics instructions also have high recall, although the number of actual positive cases (true positives and false negatives) is not significantly high enough. One other thing to note is that diuretics has a lot of false positives. Considering how rarely this instruction appears in the corpus, it might be better off not including this instruction in the grammar, if it means clinicians don't often include it anyway.

When prompting for the reasoning for the LLM as to why this specific patient data shows that instructions have to be included or not, the reasoning the LLM gives seems to be arbitrary, not really having a clear consistent explanation across different patient data. This suggests that either the MIMIC-IV dataset is missing some other data containing the information that can explain whether the instruction is included (such as other free-text notes like in MIMIC-III), or that the dataset is very noisy and not reliable enough that it can be used as a gold standard for what should be generated.

4.1.1 Front-end implementation

An end-to-end system was implemented to give clinicians a UI they may use in their discharge instruction workflow. Figures 4.3, 4.4, and 4.5 show some screenshots of the front-end implementation, which is available publicly at `https://github.com/Steve-120/discharge-dsl`. To start the system, the user specifies the patient data by giving their `subject_id`, `hadm_id`, and the directory where all patient data resides in their system. The directory contains patient data in MIMIC-IV format, split into subdirectories per patient by their `subject_id`, as well as global tables such as item ID definition tables. To generate the draft discharge instructions, a local LLM must be hosted at `localhost:8081`, accepting HTTP requests in OpenAI format. Afterwards, it generates the DischargeDSL fields from the DSPy modules, and also logical fields via a deterministic program. The evidences for the DischargeDSL field values are displayed at the top for the user to review if it is reasonable, as in Figure 4.3. These evidences have a keyboard shortcut of `F4` and `Shift+F4` for quick navigation.

The model outputs are treated as null if the confidence levels are medium or low, and

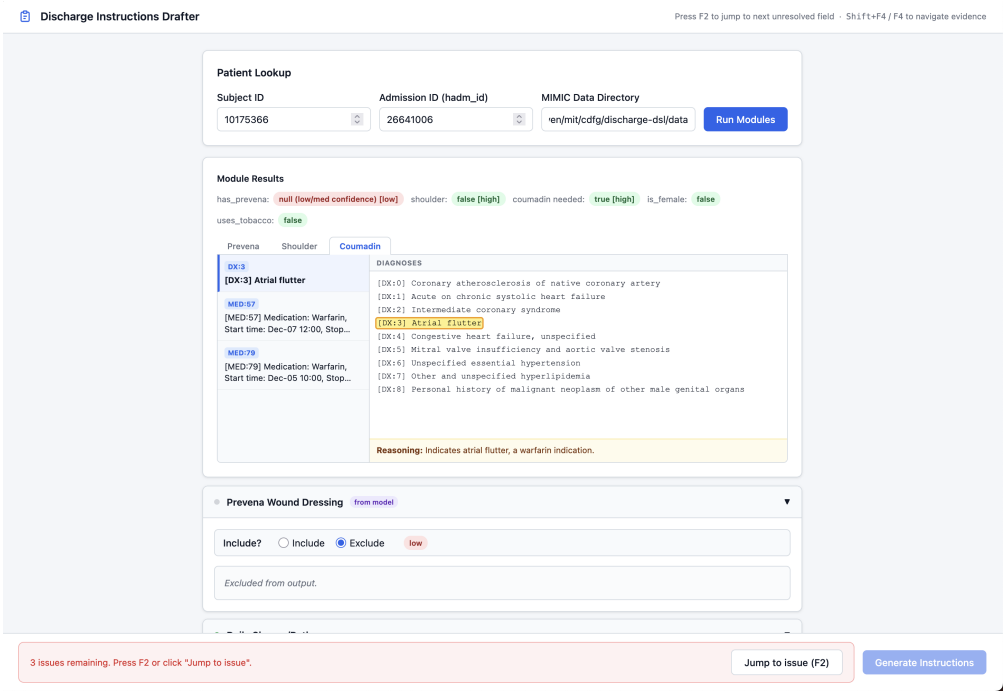

Figure 4.3: A sample patient's data loaded into the end-to-end system after finishing model prompting and ready for clinician review. This screen shows the UI for referencing the citations used by the LLM in determining its answers for the Prevena, shoulder movement, and coumadin topics.

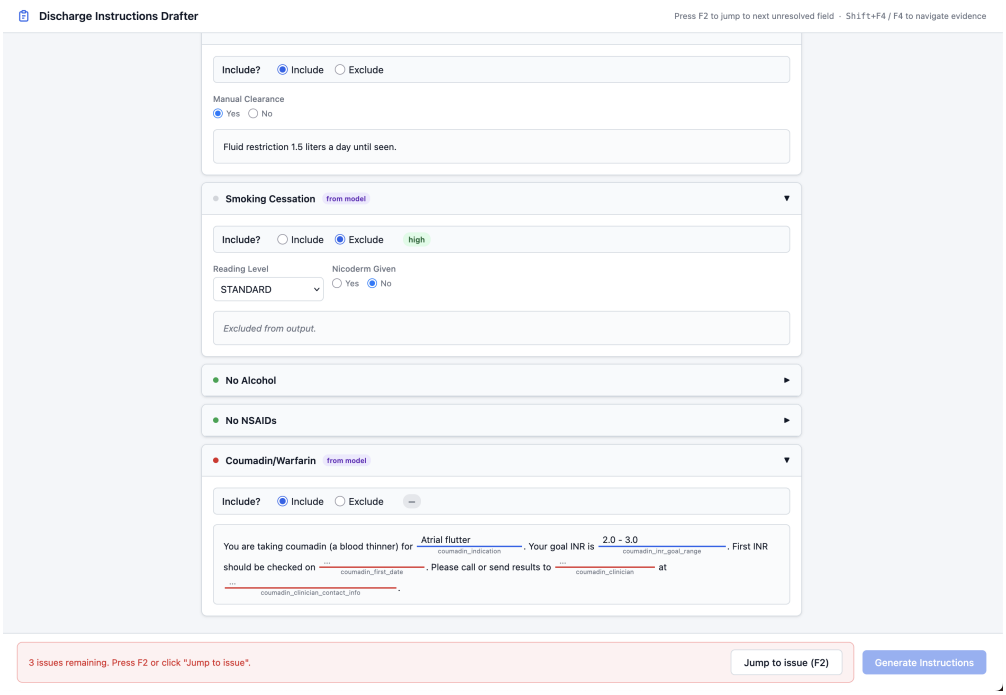

Figure 4.4: A screenshot of the UI showing the required fields the clinicians need to fill in and address, before being able to generate the finalized discharge instructions.

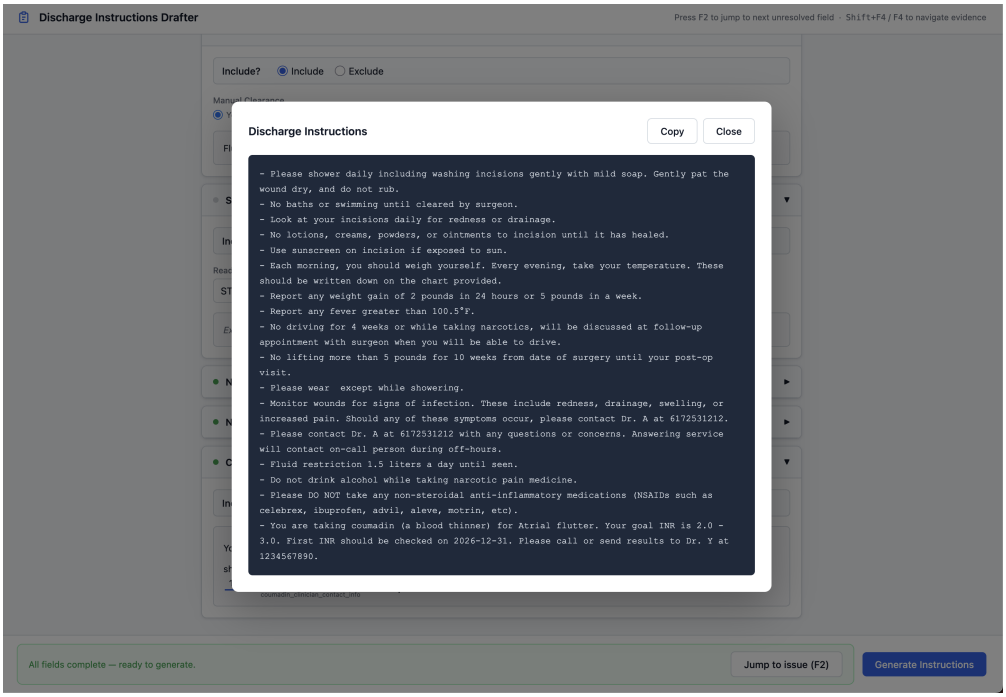

Figure 4.5: A screenshot of the UI showing the finalized discharge instructions, which can be copied to clipboard and used in the clinician's existing workflow.

the discharge instruction lines that take these fields as input are left blank. Otherwise, they are auto-filled in by the program, such as the coumadin indication and INR goal range in Figure 4.4. Any unfilled choices are flagged for review, preventing the user from generating the final instructions until they are all resolved. These flags can be quickly navigated by pressing `F2`, similar to how clinicians use existing medical systems.

Other than flags that require addressing, there are other parameters clinicians can use, such as sentence inclusion, reading levels, and clearing condition, in order to tweak specific patient outputs. Once everything is addressed, the user can generate the final output, as in Figure 4.5, which can be copied to clipboard.

4.2 Discharge Summaries

After pre-processing MIMIC-III data, there are 4991 CABG patients with discharge summaries, with 5936 discharge summary notes in total. Patients usually have multiple discharge summary

notes, due to one or more of them being an addendum to the first note.

The prototype grammar is presented in full in Listing B.3. The common sentence formats of each section are recorded in the grammar, such as `presentationSentence` and `priorCardiacHistorySentence` for the history of present illness section, `tobaccoLine` and `drugsLine` for the social history section, `cabgProcedure` and `valveProcedure` for the major surgical or invasive procedure section, etc., with the `narrativeSentence` grammar symbol serving as the catch-all symbol for all of them.

After generating a prototype grammar, standardizing into JSON files, and splitting by deterministic and nondeterministic fields, the only fields that can be deterministically filled in according to the LLM are the following:

1. Patient gender
2. Date of birth
3. Admission date
4. Discharge date
5. Service performed
6. Discharge disposition

This leaves virtually the entire rest of the discharge summary requiring LLM generation, and has to be part of the DischargeDSL grammar. Upon closer inspection of the standardized JSON files of actual discharge summaries, their free-text sections are composed of mostly the `narrativeSentence` catch-all grammar symbol rather than those of common sentence formats. This isn't desired, because if we were to try to simulate generating our own discharge summaries for these patients, this tells us that most of their discharge summary generation is performed in the refinement step after finalizing the common sentences, of which there are not that many. Thus, it is not that different from naive discharge summary generation. It is possible that the standardization of existing discharge summaries can be skipped, and that using the common sentence formats to generate new discharge summaries can still be useful.

Unfortunately, due to lack of time before the thesis deadline, the generation of DischargeDSL programs and their evaluation against actual outputs are unable to be performed in time.

4.3 Conclusion

This thesis provides a novel exploration of a semi-automated data-driven approach for developing a standardized grammar for discharge instructions and summaries, to be able to generate them using LLMs in a way where hallucinations are reduced. Specifically, a deterministic compiler from the intermediate DischargeDSL discharge instruction grammar to the finalized output is implemented, as well as a UI for clinicians to access this workflow, which is available at https://github.com/Steve-120/discharge-dsl. The implementation shown here also shows how personalization can be achieved, by targeting different languages in the template files. The framework for developing and generating DischargeDSL, such as the one-by-one pipeline for validating grammar, the ability to check the LLM's citations against the original data, and the method to curate datasets for DSPy prompt optimization, can be extended to larger scopes as well. Furthermore, the MIMIC-IV discharge instructions were also shown to be quite noisy, so that their basis as a gold standard can be questioned.

More exploration is needed with different medical datasets, such as MIMIC-III's free-text notes, to see whether such a grammar development approach is applicable in extending to LLM generation to entire discharge summaries. Specifically, the next steps are to develop a pipeline to generate a draft discharge summary from common sentences, refine it successively with free-text notes, and evaluate it against the actual discharge summary output. A severe limitation of this work is the lack of clinician review of LLM-generated outputs. This requires a user interface to show clinicians the grammar-generated text and the citations it's drawn from, to allow for selection of any additional text to be included and for correction of any errors, so this feedback can be collected to potentially improve the system and to evaluate for any time savings, accuracy metrics, and willingness to use such a system by clinicians.

Appendix A

LLM Prompts

This appendix lists the prompts used in LLM calls.

A.1 Prototype Output Grammar Development

Listing A.1: Prompt used for making prototype grammar topic lists

```
INSTRUCTIONS:

Your task is to read through the given passages of hospital discharge
instructions, and output a list of the sentences most commonly found in all
passages, modulo similar expressed meanings. Order the sentences by most often to
least, and make sure to note down absolutely all kinds of sentences found, no
matter how infrequent they're mentioned. Secondarily, explain what kinds of
variations occur in similar sentences, for each output list item. Only output in
plaintext with no extra formatting besides numbered and bulleted lists.

---
TEXT PASSAGES:

{discharge_instructions}

---
```

A.2 One-by-one Grammar Validation

Listing A.2: Prompt used for standardizing discharge instructions from the original

```
You are given a patient's discharge instructions, and a grammar spec of the
discharge instructions. Your task is to transform the free-form patient discharge
text into a standardized format using the provided grammar spec. Please output
only the DSL program specified by the grammar.

The program does not have to follow the exact format of the free-form text, but
they have to be equivalent in meaning (no extra or removed ideas). For each input,
determine whether each rule applies (see attached text). Then generate the final
formatted output. Do not skip the determination step.
```

Listing A.3: Prompt to instruct LLM not to be too strict with semantic preservation

```
Rule 1: Preserve modifiers in instructions
- Any word or phrase that modifies a verb or a noun in the free-form instruction
(e.g. until follow up with surgeon, for about one month, etc.)
 must have its meaning preserved in the final DSL line.
- Prioritize finding a modifier in the grammar closely matching the meaning in the
free-form instruction over omitting it.

Example input:
No driving for approximately one month until follow up with surgeon

✓Best output:
No driving for one month, will be discussed at follow-up appointment with surgeon.
```

Appendix B

Grammar Specifications

This appendix lists the text data involved in developing the output gramamrs for discharge instructions and discharge summaries.

Listing B.1: List of discharge summary sentence topics

```
1. Shower—and—incision—care instructions
— "Please shower daily including washing incisions gently with mild soap, no baths or swimming, and look at your incisions ."
— "Shower daily – wash incisions gently with mild soap, no baths or swimming, look at your incisions daily ."
— "Please shower daily – wash incisions gently with mild soap, no baths or swimming, look at your incisions daily ."
— "Please shower daily including washing incisions gently with mild soap, no baths or swimming until cleared by surgeon. Look at your incisions daily for redness or drainage."
— "Shower daily, no baths or swimming"
— Rarely: "Shower daily, let water flow over wounds, pat dry with a towel."
— Rarely: "Shower daily including washing incisions gently with mild soap, rinse and pat dry. DO NOT RUB"
— Rarely: "Please shower daily —wash incisions gently with mild soap, no baths or swimming, look at your incisions daily once Prevena is removed."

2. Do—not—apply topical products to the incision
— "Please NO lotions, cream, powder, or ointments to incisions ."
— "No lotions, creams, or powders to incision until it has healed."
— "Please – NO lotion, cream, powder or ointment to incisions ."

3. Daily weight—and—temperature log
— "Each morning you should weigh yourself and then in the evening take your temperature, these should be written down on the chart."
— "Each morning you should weigh yourself and then in the evening take your temperature, these should be written down on the chart provided."
— "Daily weights: keep a log."
— "Please weigh yourself at the same time every day in the morning with the same amount of
```

close on."
— "Your weight at discharge is 240.7 lbs. Please weigh yourself today at home and use this as your new baseline."

4. Driving restriction
— "No driving for approximately one month and while taking narcotics, will be discussed at follow—up appointment with surgeon when you will be able to drive."
— "No driving for one month or while taking narcotics. Driving will be discussed at follow—up appointment with surgeon."
— "No driving for approximately one month until follow up with surgeon."
— "no driving for 4 weeks and off all narcotics."
— Rarely: "No driving for 6 weeks or while taking narcotics for pain."
— One instance: "no driving for 2 weeks or until off all narcotics"
— One instance: "You may drive after 1 week as long as you are not driving while using narcotic pain medication."

5. Lifting restriction
— "No lifting more than 10 pounds for 10 weeks."
— "No lifting greater then 10 lb for 10 weeks from date of surgery."
— Rarely: "No lifting greater than 10 pounds for 6 weeks"
— Rarely: "Do not lift more than 10 lbs. for 2 months."
— One instance: "You should not push, pull, lift or carry anything heavier than 5 pounds for the next 2 weeks."
— One instance: "No excessive head turning, lifting, pushing or pulling (greater than 5 lbs) until your post op visit "

6. Contact—the—office for questions
— "Please call with any questions or concerns ___."
— "Please call the cardiac surgery office with any questions or concerns ___. Answering service will contact on—call person during off—hours."

7. Female—specific incision —care (bra)
— "Females: Please wear bra to reduce pulling on incision, avoid rubbing on lower edge."

8. Monitoring instruction
— "Monitor wounds for signs of infection. These include redness, drainage or increased pain. In the event that you have drainage from your sternal wound, please contact the surgeon at ___."
— "Monitor wounds for infection – redness, drainage, or increased pain."
— Rarely: "Please call the transplant clinic at ___ for fever > 101, chills, chest pain, shortness of breath, nausea, vomiting, diarrhea, constipation, inability to tolerate food, fluids or medications, yellowing of skin or eyes, increased abdominal pain, incisional redness, drainage or bleeding, dizziness or weakness, decreased urine output or dark, cloudy urine, swelling of abdomen or ankles, feeding tube becomes clogged or malpositioned or any other concerning symptoms."
— Rarely: "Call Dr. ___ ___ if experience: Fever > 101 or chills, increased shortness of breath, cough, chest pain."
— Rarely: "DANGER SIGNS: When to Call ___ You should call ___ or your local emergency number to be taken to the nearest emergency room for any emergency situation, such as: * Chest pain not related to your incision or angina pain, similar to the pain you had prior to surgery * Extreme shortness or breath or difficulty breathing * Severe bleeding, especially if you are on warfarin (Coumadin) * Fainting, severe lightheadedness or changes in mental status"

9. Fever—monitoring instruction
— "Report any fever greater than 100.5."
— "Report any fever greater than 101."

10. Weight—gain alert
— "Report any weight gain of 2 pounds in 24 hours or 5 pounds in 1 week."
— "Report any weight gain of greater than 2 pounds in 24 hours or 5 pounds in a week."
— Sometimes: "Daily weights: keep a log. call MD if weight goes up more than 3 lbs in 24 hours or 5 lbs over 5 days."

11. Shoulder range—of—motion encouragement
— "Encourage full shoulder range of motion, unless otherwise specified ."

12. Prevena dressing instructions (less universal but appears in many passages)
— "The Prevena wound dressing should be left on for a total of 7 days post—operatively. The date of Day # 7 should be written on a piece of tape on the canister to ensure that the nurse from the ___ or ___ facility knows when to remove the dressing and inspect the incision . If the date is not written, please alert your nurse prior to discharge."
— "You may shower, but avoid getting the dressing and suction canister soiled or saturated."
— "You will be sent home with a shower bag to hold the suction canister while bathing."
— "Once the Prevena dressing is removed, wash the incision daily with a plain white bar soap (e.g., Dove)."
— "Do not apply creams, lotions or powders to the incision after the Prevena dressing is removed."
— "If the dressing becomes soiled or saturated, turn the power off, remove the dressing and discard the unit; notify your nurse."
— "If you notice any redness, swelling or drainage, please contact your surgeon's office at ___."

13. Sometimes: Use sunscreen on incision
— "You may shower and wash incision. Gently pat the wound dry. Please shower daily . No bathing or swimming for 1 month. Use sunscreen on incision if exposed to sun."

14. Sometimes: Coumadin/warfarin
— "take 5mg of Coumadin (2 tablets) on ___ and ___, then as directed by Dr. ___"
— "Coumadin to be followed by ___ clinic at the office of Dr. ___ phone (___) fax (___)."
— "You are taking coumadin (a blood thinner) for atrial fibrillation . You goal INR is 2.0—2.5. You coumadin dosing will be managed by Dr. ___ need an appointment for blood draw (___) when discharged from ___ for coumadin management. ___"
— "Labs: ___ for Coumadin indication mechanical AVR. Goal INR 2.5—3.0. First draw ___. Results to Dr ___ ___ fax ___"
— "Monitor ___ every Mond, ___, and ___. Results should be called or faxed to Dr. ___. Warfarin should be adjusted for goal INR between 2.0 — 3.0. Dr. ___ will adjust ___ as an outpatient."

15. Rarely: Fluid restriction
— "Fluid restriction 1.5 Liters a day while on diuretics "
— "Fluid Restriction 1.5 Liters until seen"

16. Rarely: Stop smoking
— "Please stop smoking. Information was given to you on admission regarding smoking cessation."

— "Please stop smoking. Information was given to you on admission regarding smoking cessation, you have been started on nicoderm patch to assist in quiting smoking"
— "If you smoke, it is very important for you to stop. Research has shown that smoking makes vascular disease worse. Talk to your primary care physician about ways to quit smoking.

The ___ Smokers' Helpline is a FREE and confidential way to get support and information to help you quit smoking. Call ___"

17. Rarely: Don't drink alcohol while on pain medicine
— "Do not drink alcohol while taking narcotic pain medicine."
— "Alcohol use: Explained interaction with warfarin, potential increased bleeding risk associated with alcohol intake, importance of communicating changes in alcohol intake pattern with ACMS providers."
— "You have been given medications for pain control. Please do not drive, operate heavy machinery, or drink alcohol while taking these medications. As your pain decreases, take fewer ..."

18. Rarely: Don't take NSAIDs
— Only once: "Please DO NOT take any non—steroidal anti—inflammatory medications (NSAIDs such as celebrex, ibuprofen, advil, aleve, motrin, etc)."
— However, many other instructions say to take ibuprofen

19. Rarely: Wear a binder
— Only once: "Wear sternal binder at all times except for showering."
— Only once: "Chest binder at all times; please try to keep as tight as possible without compromising respiratory status."
— Only once: "Females: Please wear bra to reduce pulling on incision, avoid rubbing on lower edge, wear abdominal binder at all times."

20. Rarely: Hug a pillow when sneezing or coughing
— Only once: "Hug pillow across chest when coughing, sneezing"

21. Rarely: Check finger stick glucose
— Only once: "Monitor finger stick blood sugars twice daily. If they are consistently high Restart Metformin 500 mg twice daily."
— Only once: "Diabetes: monitor finger stick blood sugars twice daily and follow—up with your endocrinologist for further management."
— Only once: "Monitor finger stick ___ sugars twice daily. Keep a log and follow—up with your PCP"

22. Rarely: Medications (most other notes put this in a separate section)
— Rarely: Take cipro antibiotics
> "Maintain foley catheter until follow up with urology — please note needs to start cipro 500 mg twice a day — one day prior to appointment"
> Please continue Cipro course x14 days from ___ (thru ___"
> Complete 2 week course of intravenous Vancomycin and oral ciprofloxacin started on ___."

— Rarely: Take Lasix and potassium
> Lasix and potassium daily for 7 days then re—evaluate. Monitor and replete electrolytes while on lasix."
> Daily weights: take lasix 40 mg twice daily until weight down ___ then 40 mg daily. Take potassium 20 mEq once daily while taking lasix twice daily. Stop when taking lasix once daily"

> Take lasix and potassium once daily for 1 week and then stop."
> Take lasix and potassium once daily for 5 days then stop."
> Take lasix and potassium once daily for 10 days."
> Resume your preop lasix (40mg daily in AM)"

– Rarely: Take amiodarone
> "Take amiodarone 400mg twice daily for 2 more days. Starting ___, take 400mg once daily for 7 days and then (Starting ___ to 200mg once daily therafter until seen by Dr. ___."
> "Amiodarone 400mg daily for 1 week, then decrease to 200mg daily on ___."

– Rarely: Insulin
> "Please continue to monitor blood glucose, and follow up with ___ for adjustments in insulin doses"
> "NEW INSULIN PUMP SETTINGS PROGRAMMED IN YOUR PUMP:..."
> "Diabetes: Continue Metformin 1000 mg twice daily. U500 insulin 6 Units on your U100 syringe (equals to 30 units). Limit your fruit juice intake."

Listing B.2: Finalized discharge summary output grammar based on MIMIC-IV data

```
dischargeDocument = "Discharge Instructions:\n" instructionBlock ;

instructionBlock = { "- " instruction "\n" } ;
instruction      = prevenaInstructionLines
                 | showerInstructionLine
                 | bathProhibitionLine
                 | incisionInstructionLine
                 | sunscreenInstructionLine
                 | lotionProhibitionLine
                 | chartInstructionLine
                 | weightGainInstructionLine
                 | feverInstructionLine
                 | woundMonitoringLine
                 | redFlagInstructionLine
                 | drivingRestrictionLine
                 | liftingRestrictionLine
                 | contactOfficeLine
                 | femaleSpecificLine
                 | shoulderROMInstructionLine
                 | fluidRestrictionInstructionLine
                 | smokingInstructionLine
                 | alcoholInstructionLine
                 | nsaidInstructionLine
                 | binderInstructionLine
                 | otherInstructionLine

showerInstructionLine       = ( "Shower" | "Bed bath" ) " daily"
                              [ " including washing incisions"
                                [ " gently with mild soap" ]
                                [ ", sternally after Prevena is removed" ] ]
                              "." ;
                              [ "Gently pat the wound dry." ]
                              [ "DO NOT RUB." ] ;

bathProhibitionLine         = "No baths" [ " or swimming" ]
                              [ ( " until cleared by surgeon" | " for one month"
                              ) ] "." ;

incisionInstructionLine     = "Look at your incisions" [ "daily"
                              [ ( " for redness or drainage" | " once Prevena is
                              removed" ) ] ] "." ;

lotionProhibitionLine       = "NO lotions, creams, powders, or ointments to
incision"
                              [ " until it has healed" ]
                              "." ;

sunscreenInstructionLine    = "Use sunscreen on incision if exposed to sun." ;

chartInstructionLine        = weightLogging
```

```
                              [ temperatureLogging ]
                              [ " These should be written down on the chart
                              provided." ] ;

weightLogging                 = "Each morning" [ ", at the same time each day" ]
                              ", you should weigh yourself"
                              [ " with the same amount of clothes on" ]
                              "." ;

temperatureLogging            = " Every evening, take your temperature." ;

weightGainInstructionLine     = "Report any weight gain of " ( "2" | "3" ) " pounds
in 24 hours"
                              " or 5 pounds in" ( "a week" | "5 days" ) "." ;

feverInstructionLine          = "Report any fever greater than " ( "100.5" | "101" )
"." ;

drivingRestrictionLine        = "No driving " [ "or operating heavy machinery " ]
"for "
                              ( "one month" | "4 weeks" | "6 weeks" | "2 weeks" |
                              "1 week" )
                              [ " and while taking narcotics" ]
                              [ ", will be discussed at follow-up appointment
                              with surgeon"
                                  [ " when you will be able to drive" ] ]
                              "." ;

liftingRestrictionLine        =  "No " [ "pushing, pulling, carrying, or " ]
"lifting more than "
                              ( "10" | "5" ) "pounds"
                              " for " ( "10 weeks" | "6 weeks" | "2 months" | "2
                              weeks" )
                              [ " from date of surgery" ]
                              [ " until your post op visit" ]
                              "." ;

contactOfficeLine             = "Please call"
                              [ " the cardiac surgery office" ]
                              [ ( " with any questions or concerns ___." | " if
                              there is concern for wound infection." ) ]
                              [ " Answering service will contact on-call person
                              during off-hours." ] ;

femaleSpecificLine            = "Females: Please wear " ( "bra" | "abdominal binder
to chest" )
                              " to reduce pulling on incision"
                              [ ", avoid rubbing on lower edge" ]
                              [ ", bra caused skin irritation" ]
                              "." ;

woundMonitoringLine           = "Monitor wounds for signs of infection. These
```

```
include redness, drainage or increased pain."
                              [ " In the event that you have drainage from your
                              sternal wound, please call "
                                  [ ( "the surgeon at " | "the transplant clinic
                                  at " ) ]
                                  "___." ] ;

redFlagInstructionLine      = "Call " ( "the surgeon at " | "the transplant clinic
at " | "Dr. ___ at " ) "___"
                              [ " or your local emergency number" ]
                              " if you experience any of: "
                              redFlagList "." ;

redFlagList                 = redFlagItem { ", " redFlagItem } "or any other
concerning symptoms" ;

redFlagItem                 = "chills"
                            | "chest pain"
                            | "shortness of breath"
                            | "nausea"
                            | "vomiting"
                            | "diarrhea"
                            | "constipation"
                            | "inability to tolerate food, fluids, or medications"
                            | "yellowing of skin or eyes"
                            | "increased abdominal pain"
                            | "fainting"
                            | "lightheadedness"
                            | "dizziness or weakness"
                            | "decreased urine output"
                            | "dark, cloudy urine"
                            | "swelling of abdomen, ankles"
                            | "feeding tube becoming clogged or malpositioned"
                            | "severe bleeding, especially if you are on warfarin
                            (Coumadin)"

shoulderROMInstructionLine  = "Encourage full shoulder range of motion, unless
otherwise specified." ;

prevenaInstructionLines     = "Prevena instructions:"
                              prevenaDressingInstructionLine
                              prevenaShowerInstructionLine
                              prevenaShowerBagInstructionLine
                              prevenaSoiledInstructionLine
                              prevenaRemovedSoapInstructionLine
                              prevenaRemovedLotionProhibitionLine
                              prevenaRemovedMonitoringInstructionLine ;

prevenaDressingInstructionLine      = "  > The Prevena wound dressing should be
left on for a total of 7 days "
                                      "post-operatively to receive the full
                                      benefit of the therapy. The date "
```

```
                                    "of Day # 7 should be written on a piece of
                                    tape on the canister to ensure "
                                    "that the nurse from the ___ or ___ facility
                                    knows when to remove the dressing "
                                    "and inspect the incision. If the date is
                                    not written, please alert your "
                                    "nurse prior to discharge." ;

prevenaShowerInstructionLine        = "  > You may shower, but avoid getting the
dressing and suction canister soiled "
                                    "or saturated." ;

prevenaShowerBagInstructionLine     = "  > You will be sent home with a shower bag
to hold the suction canister while "
                                    "bathing." ;

prevenaSoiledInstructionLine        = "  > If the dressing becomes soiled or
saturated, turn the power off, remove the"
                                    " dressing, and discard the unit. Should
                                    this happen, please notify your ___ "
                                    "nurse, so they may make plans to see you
                                    the following day to assess your "
                                    "incision." ;

prevenaRemovedSoapInstructionLine   = "  > Once the Prevena dressing is removed,
you may wash your incision daily with"
                                    " a plain white bar soap, such as Dove." ;

prevenaRemovedLotionProhibitionLine = "  > Do not apply any creams, lotions, or
powders to your incision and monitor "
                                    "it daily." ;

prevenaRemovedMonitoringInstructionLine = " > If you notice any redness, swelling
or drainage, please contact your "
                                    "surgeon's office at ___." ;

fluidRestrictionInstructionLine = "Fluid restriction 1.5 liters a day " ( "until
seen" | "while on diuretics" ) "." ;

smokingInstructionLine      = "Please stop smoking."
                              [ " Information was given to you on admission
                              regarding smoking cessation." ]
                              [ " You have been started on nicoderm patch to
                              assist in quitting smoking." ]
                              [ " If you smoke, it is very important for you to
                              stop. Research has shown that "
                                  "smoking makes vascular disease worse. Talk to
                                  your primary care physician "
                                  "about ways to quit smoking. The ___ Smokers'
                                  Helpline is a FREE and confidential "
                                  "way to get support and information to help you
                                  quit smoking. Call ___." ] ;
```

```
alcoholInstructionLine      = "Do not drink alcohol while taking narcotic pain
medicine." ;

nsaidInstructionLine        = "Please DO NOT take any non-steroidal
anti-inflammatory medications (NSAIDs such as "
                              "celebrex, ibuprofen, advil, aleve, motrin, etc)." ;

binderInstructionLine       = "Wear " ( "sternal" | "chest" | "abdominal" ) "
binder at all times"
                              [ " except for showering" ] "."
                              [ " Please try to keep as tight as possible without
                              compromising respiratory status." ] ;

otherInstructionLine        = ( WORD | NUMBER ) { ( " " ) ( WORD | NUMBER ) } [
"." | "!" | "?" ] ;

NUMBER         = DIGIT+ [ "." DIGIT+ ] ;
DIGIT          = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9" ;
WORD           = LETTER { LETTER | "-" | "'" } ;
LETTER         = "A".."Z" | "a".."z" ;
```

Listing B.3: Prototype discharge summary output grammar based on MIMIC-III data

```
grammar CABGDischargeSummary;

/* ================================================================
 *                        PARSER RULES
 * ================================================================ */

document
    : dischargeSummary+ EOF
    ;

dischargeSummary
    : BEGIN_DELIMITER
      headerBlock
      bodySection*
      signatureBlock?
      END_DELIMITER
    ;

/* -------- Header block (always at top of every summary) -------- */

headerBlock
    : admissionLine
      dobSexLine?
      serviceLine?
      attendingLine?
    ;

admissionLine
    : 'Admission' 'Date' ':' dateValue 'Discharge' 'Date' ':' dateValue
    ;

dobSexLine
    : 'Date' 'of' 'Birth' ':' dateValue 'Sex' ':' sexValue
    ;

serviceLine
    : 'Service' ':' serviceValue?
    ;

attendingLine
    : 'Attending' ':' DEID
    ;

sexValue
    : ('M' | 'F')
    ;

serviceValue
    : ('CARDIOTHORACIC'
      | 'Cardiothoracic' 'Surgery'
```

```
      | 'CARDIOTHORACIC' 'SURGERY'
      | 'CARDTHOR'
      | 'SURGERY'
      | textPhrase)
    ;

dateValue
    : DEID          /* e.g. [**2190-5-16**] */
    ;

/* -------- All body sections, in any order observed -------- */

bodySection
    : allergiesSection
    | chiefComplaintSection
    | majorSurgicalProcedureSection
    | hpiSection
    | pastMedicalHistorySection
    | pastSurgicalHistorySection
    | socialHistorySection
    | familyHistorySection
    | reviewOfSystemsSection
    | medicationsOnAdmissionSection
    | preoperativeMedicationsSection
    | physicalExamSection
    | pertinentResultsSection
    | hospitalCourseSection
    | conditionAtDischargeSection
    | dischargeStatusSection
    | dischargeDispositionSection
    | facilitySection
    | dischargeDiagnosisSection
    | secondaryDiagnosesSection
    | dischargeMedicationsSection
    | dischargeConditionSection
    | dischargeInstructionsSection
    | followupInstructionsSection
    | dietSection
    | treatmentsSection
    | impressionSection
    | laboratoryDataSection
    | dictationFooterSection
    | unknownSection      /* fallback for unseen but well-formed sections */
    ;

/* ====================================================================
 *  HISTORY OF PRESENT ILLNESS
 * ==================================================================== */

hpiSection
    : hpiHeader hpiSentence+
    ;
```

```
hpiHeader
    : ('HISTORY' 'OF' 'PRESENT' 'ILLNESS' ':'
      | 'HISTORY' 'OF' 'THE' 'PRESENT' 'ILLNESS' ':'
      | 'History' 'of' 'Present' 'Illness' ':')
    ;

hpiSentence
    : presentationSentence
    | priorCardiacHistorySentence
    | symptomProgressionSentence
    | transferSentence
    | diagnosticDiscoverySentence
    | surgicalReferralSentence
    | narrativeSentence
    ;

presentationSentence
    : ageSexIntro? historyClause?
      ('presented' 'with' | 'was' 'admitted' 'with' | 'reported'
        | 'admitted' 'with' | 'was' 'admitted' 'to' admittedLocation 'with')
      symptomList '.'
    ;

ageSexIntro
    : ('This' | 'The' 'patient' 'is')? 'a'? AGE ('year' '-'? 'old' | 'yo')
      ('male' | 'female' | 'man' | 'woman' | 'gentleman' | 'lady'
       | 'white' 'male' | 'white' 'female')
    ;

historyClause
    : ('with' ('history' 'of' | 'a' 'history' 'of' | 'known')
       diagnosisList ',')
    ;

symptomProgressionSentence
    : ('Symptoms' | 'The' 'patient' 'noted' | 'He' | 'She')
      ('progressed' | 'worsened' | 'increased' | 'continued')
      textPhrase '.'
    ;

priorCardiacHistorySentence
    : ('Patient' 'is' 'status' 'post' | 's/p' | 'Status' 'post')
      procedureList '.'
    ;

transferSentence
    : ('She' | 'He' | 'The' 'patient' | 'Patient')
      ('was' 'transferred' | 'was' 'referred' | 'was' 'sent')
      ('to' | 'from')
      facilityReference
      ('for' textPhrase)? '.'
```

```
    ;

diagnosticDiscoverySentence
    : testName
      ('revealed' | 'showed' | 'demonstrated' | 'found' | 'was' 'remarkable' 'for')
      findingList '.'
    ;

surgicalReferralSentence
    : ('She' | 'He' | 'The' 'patient' | 'Patient')
      'was' 'referred'
      ('to' ('cardiothoracic' 'surgery' | 'cardiac' 'surgery'
            | 'surgery' | 'cardiac' 'surgery' 'for' textPhrase))
      '.'
    ;

admittedLocation
    : ('an' 'outside' 'hospital'
      | 'the' 'ED'
      | 'the' 'emergency' 'department'
      | facilityReference)
    ;

facilityReference
    : (DEID | textPhrase) ('Hospital' DEID?)?
    ;

testName
    : ('Cardiac' 'catheterization'
      | 'cardiac' 'catheterization'
      | 'Cardiac' 'cath'
      | 'Stress' 'test'
      | 'Echocardiogram'
      | 'echocardiogram'
      | 'EKG'
      | 'ECG'
      | 'CXR'
      | 'CT' 'scan'
      | 'MRI'
      | 'Carotid' 'ultrasound'
      | 'Exercise' 'tolerance' 'test'
      | 'SPECT'
      | 'Transthoracic' 'echocardiogram'
      | 'LHC')
    ;

findingList
    : finding (',' finding)* (('and') finding)?
    ;

finding
    : (PERCENT? 'lesion' anatomicalSite?
```

```
      | PERCENT? 'stenosis' anatomicalSite?
      | 'ejection' 'fraction' ('of' NUMBER PERCENT?)?
      | 'EF' NUMBER PERCENT?
      | textPhrase)
    ;

anatomicalSite
    : ('of' 'the')?
      ('left' 'anterior' 'descending' ('coronary' 'artery')?
      | 'right' 'coronary' 'artery'
      | 'circumflex'
      | 'diagonal' NUMBER?
      | 'obtuse' 'marginal' NUMBER?
      | 'ramus'
      | 'left' 'main'
      | 'PDA'
      | 'OM' NUMBER?
      | 'LAD'
      | 'RCA'
      | 'LCX'
      | 'LMCA'
      | textPhrase)
    ;

/* Catch-all narrative sentence: matches any remaining sentence-like
   text inside the HPI section. */
narrativeSentence
    : sentenceText '.'
    ;

/* ========================================================================
 *  PAST MEDICAL HISTORY / PAST SURGICAL HISTORY / PMH / PSH
 * ======================================================================== */

pastMedicalHistorySection
    : pastMedicalHeader (numberedDiagnosisList | diagnosisList | textBlock)
    ;

pastMedicalHeader
    : ('PAST' 'MEDICAL' 'HISTORY' ':'
      | 'Past' 'Medical' 'History' ':'
      | 'PMH' ':')
    ;

pastSurgicalHistorySection
    : pastSurgicalHeader (numberedDiagnosisList | diagnosisList | textBlock)
    ;

pastSurgicalHeader
    : ('PAST' 'SURGICAL' 'HISTORY' ':'
      | 'Past' 'Surgical' 'History' ':'
      | 'PSH' ':')
```

```
    ;

numberedDiagnosisList
    : numberedItem+
    ;

numberedItem
    : NUMBER '.' diagnosisItem '.'?
    ;

diagnosisList
    : diagnosisItem (',' diagnosisItem)* (('and') diagnosisItem)? '.'?
    ;

diagnosisItem
    : (statusPostPrefix)? textPhrase
    ;

statusPostPrefix
    : ('Status' 'post' | 'status' 'post' | 's/p' | 'h/o' | 'History' 'of')
    ;

procedureList
    : procedureItem (',' procedureItem)* (('and') procedureItem)?
    ;

procedureItem
    : textPhrase
    ;

/* ================================================================
 *  SOCIAL HISTORY / FAMILY HISTORY / REVIEW OF SYSTEMS
 * ================================================================ */

socialHistorySection
    : socialHistoryHeader socialHistoryContent
    ;

socialHistoryHeader
    : ('SOCIAL' 'HISTORY' ':' | 'Social' 'History' ':')
    ;

socialHistoryContent
    : socialHistoryItem+
    ;

socialHistoryItem
    : tobaccoLine
    | etohLine
    | drugsLine
    | livesWithLine
    | occupationLine
```

```
    | raceLine
    | narrativeSentence
    ;

tobaccoLine
    : ('Tobacco' | 'TOBACCO' | 'smokes' | 'Smokes') ':'? (textPhrase | 'none' |
    'None') '.'?
    ;

etohLine
    : ('ETOH' | 'EtOH' | 'ETOH') ':'? (textPhrase | 'None' | 'none') '.'?
    ;

drugsLine
    : ('Drugs' | 'DRUGS') ':'? (textPhrase | 'None' | 'none') '.'?
    ;

livesWithLine
    : 'Lives' 'with' ':'? textPhrase '.'?
    ;

occupationLine
    : 'Occupation' ':' textPhrase '.'?
    ;

raceLine
    : 'Race' ':' textPhrase '.'?
    ;

familyHistorySection
    : familyHistoryHeader familyHistoryContent
    ;

familyHistoryHeader
    : ('FAMILY' 'HISTORY' ':' | 'Family' 'History' ':')
    ;

familyHistoryContent
    : ('non-contributory' | 'noncontributory' | 'Non-contributory') '.'?
    | textBlock
    ;

reviewOfSystemsSection
    : reviewOfSystemsHeader textBlock
    ;

reviewOfSystemsHeader
    : ('REVIEW' 'OF' 'SYSTEMS' ':' | 'Review' 'of' 'Systems' ':' | 'ROS' ':')
    ;

/* ================================================================
 *  ALLERGIES
```

```
 * ========================================================================= */

allergiesSection
    : allergiesHeader allergyContent
    ;

allergiesHeader
    : ('ALLERGIES' ':' | 'Allergies' ':')
    ;

allergyContent
    : noKnownAllergies
    | allergenList
    ;

noKnownAllergies
    : ('No' 'Known' 'Allergies' ('/' 'Adverse' 'Drug' 'Reactions')?
      | 'No' 'known' 'drug' 'allergies'
      | 'NKDA') '.'?
    ;

allergenList
    : allergen (('/' | ',' | 'and') allergen)*
    ;

allergen
    : WORD+
    ;

/* ========================================================================
 *  CHIEF COMPLAINT
 * ========================================================================= */

chiefComplaintSection
    : chiefComplaintHeader chiefComplaintContent
    ;

chiefComplaintHeader
    : ('CHIEF' 'COMPLAINT' ':' | 'Chief' 'Complaint' ':')
    ;

chiefComplaintContent
    : symptomList '.'?
    ;

symptomList
    : symptom (',' symptom)* (('and') symptom)?
    ;

symptom
    : textPhrase
    ;
```

```
/* ========================================================================
 *  MAJOR SURGICAL OR INVASIVE PROCEDURE
 * ======================================================================== */

majorSurgicalProcedureSection
    : majorSurgicalProcedureHeader procedureBlock
    ;

majorSurgicalProcedureHeader
    : 'Major' 'Surgical' 'or' 'Invasive' 'Procedure' ':'
    ;

procedureBlock
    : procedureEntry+
    ;

procedureEntry
    : (dateValue)? (NUMBER '.')? procedureDescription
    ;

procedureDescription
    : cabgProcedure
    | valveProcedure
    | catheterizationProcedure
    | otherProcedure
    ;

cabgProcedure
    : ('CABG' | 'Coronary' 'artery' 'bypass' 'graft' ('ing')?
       | 'Coronary' 'bypass' 'grafting'
       | 'cabg')
      ('x' NUMBER | 'X' NUMBER | 'times' NUMBER)?
      ('(' graftList ')')?
      ('with' textPhrase)?
    ;

graftList
    : graftSpec (',' graftSpec)* (('and' | ';') graftSpec)?
    ;

graftSpec
    : (graftSource ('to' | '->' | '-' '>') graftTarget)
    | textPhrase
    ;

graftSource
    : ('LIMA' | 'left' 'internal' 'mammary' 'artery'
      | 'SVG' | 'saphenous' 'vein' 'graft'
      | 'RIMA' | 'right' 'internal' 'mammary' 'artery'
      | 'radial' 'artery'
      | textPhrase)
```

```
    ;

graftTarget
    : anatomicalSite
    | textPhrase
    ;

valveProcedure
    : ('AVR' | 'Aortic' 'valve' 'replacement'
      | 'MVR' | 'Mitral' 'valve' 'replacement'
      | 'MVRepair' | 'Mitral' 'valve' 'repair'
      | 'TVR' | 'Tricuspid' 'valve' 'replacement')
      ('with' textPhrase)?
    ;

catheterizationProcedure
    : ('Cardiac' 'cath' | 'Cardiac' 'catheterization' | 'cardiac' 'cath')
      textPhrase?
    ;

otherProcedure
    : textPhrase
    ;

/* ================================================================
 *  MEDICATIONS - ADMISSION / PREOPERATIVE / DISCHARGE
 * ================================================================ */

medicationsOnAdmissionSection
    : medicationsOnAdmissionHeader medicationList
    ;

medicationsOnAdmissionHeader
    : ('MEDICATIONS' 'ON' 'ADMISSION' ':'
      | 'Medications' 'on' 'Admission' ':'
      | 'MEDICATIONS' ':'
      | 'ADMISSION' 'MEDICATIONS' ':')
    ;

preoperativeMedicationsSection
    : preoperativeMedicationsHeader medicationList
    ;

preoperativeMedicationsHeader
    : ('PREOPERATIVE' 'MEDICATIONS' ':')
    ;

dischargeMedicationsSection
    : dischargeMedicationsHeader medicationList
    ;

dischargeMedicationsHeader
```

```
    : ('DISCHARGE' 'MEDICATIONS' ':'
      | 'Discharge' 'Medications' 'on' 'Discharge' ':'
      | 'MEDICATIONS' 'ON' 'DISCHARGE' ':'
      | 'Discharge' 'Medications' ':')
    ;

medicationList
    : medication (',' medication)*
    | numberedMedication+
    | medicationName+      /* simple comma- or newline-separated list */
    ;

numberedMedication
    : NUMBER '.' medication ('Disp' ':' dispensingInfo)?
    ;

medication
    : medicationName
      strength?
      formulation?
      sigClause?
      frequencyClause?
      durationClause?
      indicationClause?
    ;

medicationName
    : WORD ('-' WORD)*
    ;

strength
    : NUMBER ('.' NUMBER)? unit
    ;

unit
    : ('mg' | 'mL' | 'g' | 'mcg' | 'units' | 'unit' | 'mEq'
      | 'milliequivalents' | '%' | 'puffs' | 'puff' | 'cc')
    ;

formulation
    : ('Tablet' | 'Tablets'
      | 'Capsule' | 'Capsules'
      | 'Suspension'
      | 'Solution'
      | 'Suppository'
      | 'Aerosol'
      | 'Lotion'
      | 'Ointment'
      | 'Elixir'
      | 'Patch'
      | 'Extended' 'Release' ('24' 'hr')?
      | 'Delayed' 'Release' ('(' 'E' '.' 'C' '.' ')')?
```

```
      | 'Sustained' 'Release' ('24HR' | '24' 'hr')?
      | 'Chewable')
    ;

sigClause
    : 'Sig' ':' sigContent
    ;

sigContent
    : textPhrase
    ;

frequencyClause
    : ('PO' | 'IV' | 'IM' | 'SC' | 'subQ' | 'po' | 'iv'
       | 'Rectal' | 'Topical' | 'Inhalation' | 'Injection'
       | 'Sublingual')?
      frequencyCode
      ('(' textPhrase ')')?
    ;

frequencyCode
    : ('q.d.' | 'q.day' | 'qd' | 'q' 'day' | 'DAILY' | 'Daily' | 'once' 'a' 'day'
      | 'b.i.d.' | 'bid' | 'BID' | 'twice' 'a' 'day' | 'twice' 'per' 'day'
      | 't.i.d.' | 'tid' | 'TID' | 'three' 'times' 'a' 'day'
      | 'q.i.d.' | 'qid' | 'QID' | 'four' 'times' 'a' 'day'
      | 'q.h.s.' | 'qhs' | 'HS' | 'h.s.' | 'at' 'bedtime'
      | 'p.r.n.' | 'prn' | 'PRN' | 'as' 'needed'
      | 'q' NUMBER 'h' | 'Q' NUMBER 'H'
      | 'q' NUMBER 'to' NUMBER 'hours'
      | 'q.4h.' | 'q.6h.' | 'q.8h.' | 'q.12h.'
      | 'ASDIR')
    ;

durationClause
    : ('times' | 'for' | 'x')
      NUMBER ('days' | 'day' | 'weeks' | 'week' | 'doses' | 'dose' | 'month')
    ;

indicationClause
    : ('for' | 'as' 'needed' 'for') textPhrase
    ;

dispensingInfo
    : '*' NUMBER textPhrase '*' 'Refills' ':' '*' NUMBER '*'
    ;

/* ========================================================================
 *  PHYSICAL EXAM
 * ======================================================================== */

physicalExamSection
    : physicalExamHeader physicalExamBody
```

```
    ;

physicalExamHeader
    : ('PHYSICAL' 'EXAMINATION' ':'
      | 'PHYSICAL' 'EXAMINATION' 'ON' 'PRESENTATION' ':'
      | 'PHYSICAL' 'EXAMINATION' 'ON' 'DISCHARGE' ':'
      | 'DISCHARGE' 'PHYSICAL' 'EXAMINATION' ':'
      | 'PREOPERATIVE' 'PHYSICAL' 'EXAMINATION' ':'
      | 'Physical' 'Exam' ':')
    ;

physicalExamBody
    : vitalsBlock? examFinding+
    ;

vitalsBlock
    : (vsLine | vitalsTemplate)
    ;

vsLine
    : ('VS' ':' | 'Vital' 'signs' ':')? vitalEntry+
    ;

vitalsTemplate
    : (pulseEntry | respEntry | satEntry | bpEntry | heightEntry | weightEntry)+
    ;

vitalEntry
    : tempEntry
    | pulseEntry
    | bpEntry
    | respEntry
    | satEntry
    | heightEntry
    | weightEntry
    ;

tempEntry : ('T' | 'Temp' | 'Temperature') NUMBER ('.' NUMBER)? ('F' | 'C')? ;
pulseEntry : ('HR' | 'Pulse' | 'P') ':'? NUMBER ;
bpEntry   : ('BP' | 'B/P' | 'Blood' 'pressure') ':'?
            (sideValue | NUMBER '/' NUMBER) ;
respEntry : ('RR' | 'Resp' | 'Respiratory' 'rate') ':'? NUMBER ;
satEntry  : ('SpO2' | 'O2' 'sat' | 'oxygen' 'saturation') ':'? NUMBER PERCENT?
('RA' | 'on' 'room' 'air' | 'on' NUMBER 'L' textPhrase)? ;
heightEntry: 'Height' ':' textPhrase ;
weightEntry: 'Weight' ':' textPhrase ;

sideValue
    : ('Right' ':' NUMBER '/' NUMBER ('Left' ':' NUMBER '/' NUMBER)?
      | 'Left' ':' NUMBER '/' NUMBER)
    ;
```

```
examFinding
    : examSystemLine
    | narrativeSentence
    ;

examSystemLine
    : examSystemLabel ':' examSystemContent
    ;

examSystemLabel
    : ('General' | 'Gen'
      | 'HEENT'
      | 'Neck'
      | 'Chest' | 'Pulm' | 'Respiratory' | 'Lungs'
      | 'Heart' | 'CV' | 'Cv' | 'Cardiovascular'
      | 'Abdomen' | 'Abd' | 'GI' | 'Gastrointestinal'
      | 'Extremities' | 'Ext'
      | 'Neuro' | 'Neurologic'
      | 'Skin'
      | 'Pulses'
      | 'Rectal'
      | 'GU'
      | 'HEME'
      | 'ID'
      | 'DISPO'
      | 'N'              /* shorthand for Neuro */
      | 'Carotid' 'Bruit')
    ;

examSystemContent
    : textPhrase ('[' ('x' | 'xx') ']')? /* checkbox marks like [x] */
      (subFinding)*
    ;

subFinding
    : pulsesSubFinding
    | textPhrase
    ;

pulsesSubFinding
    : ('Femoral' | 'DP' | 'PT' | 'Radial')
      'Right' ':' textPhrase
      'Left' ':' textPhrase
    ;

/* ================================================================
 *  PERTINENT RESULTS / LABORATORY DATA / ECHO REPORTS
 * ================================================================ */

pertinentResultsSection
    : pertinentResultsHeader resultsBody
    ;
```

```
pertinentResultsHeader
    : ('Pertinent' 'Results' ':' | 'PERTINENT' 'RESULTS' ':')
    ;

resultsBody
    : (labResultLine | echoFindingLine | imagingReportLine | narrativeSentence)+
    ;

labResultLine
    : dateValue TIME? 'BLOOD' labAnalyteValue+
    ;

labAnalyteValue
    : analyteName '-' NUMBER ('.' NUMBER)? ('*' | '#')*
    ;

analyteName
    : WORD+
    ;

echoFindingLine
    : echoSectionLabel ':' textPhrase
    ;

echoSectionLabel
    : ('LEFT' 'ATRIUM'
      | 'RIGHT' 'ATRIUM' '/' 'INTERATRIAL' 'SEPTUM'
      | 'LEFT' 'VENTRICLE'
      | 'RIGHT' 'VENTRICLE'
      | 'LV' 'WALL' 'MOTION'
      | 'AORTA'
      | 'AORTIC' 'VALVE'
      | 'MITRAL' 'VALVE'
      | 'TRICUSPID' 'VALVE'
      | 'PULMONIC' 'VALVE'
      | 'PERICARDIUM'
      | 'GENERAL' 'COMMENTS'
      | 'IMPRESSION'
      | 'PRE-BYPASS'
      | 'POSTBYPASS'
      | 'Conclusions')
    ;

imagingReportLine
    : ('CXR' | 'CT' | 'MRI' | 'LHC' | 'LENI' | 'Carotid' 'U/S' | 'Panorex')
      dateValue? ':' textPhrase
    ;

laboratoryDataSection
    : laboratoryDataHeader textBlock
    ;
```

```
laboratoryDataHeader
    : ('LABORATORY' 'DATA' ':' | 'LABORATORY' ':')
    ;

/* ================================================================
 *  HOSPITAL COURSE / BRIEF HOSPITAL COURSE
 * ================================================================ */

hospitalCourseSection
    : hospitalCourseHeader hospitalCourseBody
    ;

hospitalCourseHeader
    : ('HOSPITAL' 'COURSE' ':'
      | 'Brief' 'Hospital' 'Course' ':'
      | 'BRIEF' 'HOSPITAL' 'COURSE' ':')
    ;

hospitalCourseBody
    : hospitalCourseSentence+
    ;

hospitalCourseSentence
    : admissionSentence
    | operativeSentence
    | postopSentence
    | complicationSentence
    | dischargeSentence
    | narrativeSentence
    ;

admissionSentence
    : ('The' 'patient' 'was' 'admitted'
       | 'Mr' '.' DEID? 'was' 'admitted'
       | 'Mrs' '.' DEID? 'was' 'admitted'
       | 'Admitted')
      ('to' textPhrase)?
      ('on' dateValue)?
      textPhrase? '.'
    ;

operativeSentence
    : ('The' 'patient' 'was' 'taken' 'to' 'the' 'Operating' 'Room'
       | 'The' 'patient' 'was' 'brought' 'to' 'the' 'operating' 'room'
       | 'Taken' 'to' 'the' 'OR')
      ('on' dateValue)?
      ('with' 'Dr' '.' DEID)?
      ('where' 'the' 'patient' 'underwent' procedureDescription)?
      '.'
    ;
```

```
postopSentence
    : ('On' 'postoperative' 'day' ('number')? (NUMBER | textPhrase) ',')
      textPhrase '.'
    ;

complicationSentence
    : ('The' 'patient' ('developed' | 'experienced' | 'was' 'noted' 'to' 'have'))
      textPhrase '.'
    ;

dischargeSentence
    : ('The' 'patient' 'was' 'discharged'
      | 'She' 'was' 'discharged'
      | 'He' 'was' 'discharged'
      | 'Discharged')
      ('to' dischargeDestination)?
      ('on' dateValue)?
      ('in' textPhrase 'condition')?
      '.'
    ;

dischargeDestination
    : ('home' | 'Home'
      | 'rehab' | 'rehabilitation' | 'Rehabilitation'
      | 'rehabilitation' 'services'
      | 'rehabilitation' 'center'
      | 'extended' 'care'
      | 'Extended' 'Care'
      | textPhrase)
    ;

/* ========================================================================
 *  CONDITION / DISPOSITION / DIAGNOSIS / INSTRUCTIONS at DISCHARGE
 * ======================================================================== */

conditionAtDischargeSection
    : conditionAtDischargeHeader textBlock
    ;

conditionAtDischargeHeader
    : ('CONDITION' 'AT' 'DISCHARGE' ':'
      | 'CONDITION' 'ON' 'DISCHARGE' ':'
      | 'DISCHARGE' 'CONDITION' ':')
    ;

dischargeStatusSection
    : dischargeStatusHeader textBlock
    ;

dischargeStatusHeader
    : 'DISCHARGE' 'STATUS' ':'
    ;
```

```
dischargeDispositionSection
    : dischargeDispositionHeader dispositionValue
    ;

dischargeDispositionHeader
    : ('Discharge' 'Disposition' ':' | 'DISCHARGE' 'DISPOSITION' ':')
    ;

dispositionValue
    : ('Home'
      | 'Home' 'With' 'Service'
      | 'Extended' 'Care'
      | 'Rehabilitation' 'Center'
      | textPhrase)
    ;

facilitySection
    : facilityHeader textPhrase
    ;

facilityHeader
    : 'Facility' ':'
    ;

dischargeDiagnosisSection
    : dischargeDiagnosisHeader diagnosisBody
    ;

dischargeDiagnosisHeader
    : ('DISCHARGE' 'DIAGNOSES' ':'
      | 'DISCHARGE' 'DIAGNOSIS' ':'
      | 'Discharge' 'Diagnosis' ':')
    ;

diagnosisBody
    : numberedDiagnosisList
    | diagnosisList
    | textBlock
    ;

secondaryDiagnosesSection
    : secondaryDiagnosesHeader (numberedDiagnosisList | diagnosisList)
    ;

secondaryDiagnosesHeader
    : 'SECONDARY' 'DIAGNOSES' ':'
    ;

dischargeConditionSection
    : dischargeConditionHeader dischargeConditionBody
    ;
```

```
dischargeConditionHeader
    : 'Discharge' 'Condition' ':'
    ;

dischargeConditionBody
    : dischargeConditionItem+
    ;

dischargeConditionItem
    : mentalStatusLine
    | levelOfConsciousnessLine
    | activityStatusLine
    | incisionsLine
    | edemaLine
    | narrativeSentence
    ;

mentalStatusLine
    : 'Mental' 'Status' ':' textPhrase '.'?
    ;

levelOfConsciousnessLine
    : 'Level' 'of' 'Consciousness' ':' textPhrase '.'?
    ;

activityStatusLine
    : 'Activity' 'Status' ':' textPhrase '.'?
    ;

incisionsLine
    : 'Incisions' ':' textPhrase '.'?
    ;

edemaLine
    : 'Edema' textPhrase '.'?
    ;

dischargeInstructionsSection
    : dischargeInstructionsHeader textBlock
    ;

dischargeInstructionsHeader
    : ('DISCHARGE' 'INSTRUCTIONS' ':'
      | 'DISCHARGE' 'INSTRUCTIONS' '/' 'FOLLOWUP' ':'
      | 'DISCHARGE' 'INSTRUCTIONS' '-' 'FOLLOWUP' ':'
      | 'Discharge' 'Instructions' ':')
    ;

followupInstructionsSection
    : followupInstructionsHeader followupBody
    ;
```

```
followupInstructionsHeader
    : ('Followup' 'Instructions' ':'
      | 'DISCHARGE' 'FOLLOWUP' ':'
      | 'Follow' '-' 'up' 'Instructions' ':')
    ;

followupBody
    : followupItem+
    ;

followupItem
    : appointmentLine
    | numberedFollowupItem
    | narrativeSentence
    ;

numberedFollowupItem
    : NUMBER '.' textPhrase '.'?
    ;

appointmentLine
    : appointmentRole ':' appointmentDetail
    ;

appointmentRole
    : ('Surgeon'
      | 'Cardiologist'
      | 'Primary' 'Care' ('Dr' '.')?
      | 'Wound' 'check'
      | 'Urologist'
      | 'Provider'
      | 'NURSE')
    ;

appointmentDetail
    : (DEID | textPhrase)
      ('Phone' ':' DEID)?
      ('Date' '/' 'Time' ':' dateValue TIME?)?
    ;

dietSection
    : dietHeader textBlock
    ;

dietHeader
    : 'DIET' ':'
    ;

treatmentsSection
    : treatmentsHeader textBlock
    ;
```

```
treatmentsHeader
    : 'TREATMENTS' ':'
    ;

impressionSection
    : impressionHeader textBlock
    ;

impressionHeader
    : ('IMPRESSION' ':' | 'Impression' ':')
    ;

/* ================================================================
 *  DICTATION FOOTER (electronic signature, dictated by, MEDQUIST36 etc.)
 * ================================================================ */

dictationFooterSection
    : (signatureLine
      | dictatedByLine
      | medquistLine
      | dictationDateLine
      | transcriptionDateLine
      | jobNumberLine
      | completedByLine)+
    ;

signatureBlock
    : dictationFooterSection
    ;

signatureLine
    : DEID+ ('M' '.' 'D' '.' | 'MD') DEID?
    ;

dictatedByLine
    : 'Dictated' 'By' ':' DEID
    ;

medquistLine
    : 'MEDQUIST36'
    ;

dictationDateLine
    : 'D' ':' dateValue TIME
    ;

transcriptionDateLine
    : 'T' ':' dateValue TIME
    ;

jobNumberLine
```

```
    : 'JOB' '#' ':' DEID
    | 'Job' '#' ':' DEID
    ;

completedByLine
    : 'Completed' 'by' ':' dateValue
    ;

/* ================================================================
 *  GENERIC FALLBACK SECTION - for any "Header: text" not enumerated
 * ================================================================ */

unknownSection
    : sectionHeader textBlock
    ;

sectionHeader
    : (WORD ('-' WORD | '/' WORD)*)+ ':'
    ;

/* ================================================================
 *  TEXT BUILDING BLOCKS
 * ================================================================ */

textBlock
    : sentenceText+
    ;

sentenceText
    : textToken+
    ;

textPhrase
    : textToken+
    ;

textToken
    : WORD
    | NUMBER
    | DEID
    | TIME
    | PERCENT
    | AGE
    | WS
    | '.'
    | ','
    | ';'
    | ':'
    | '-'
    | '/'
    | '('
    | ')'
```

```
    | '['
    | ']'
    | '+'
    | '*'
    | '#'
    | '\''
    | '"'
    | '?'
    | '>'
    | '<'
    | '='
    | '&'
    | ANY_CHAR
    ;


/* ================================================================
 *                       LEXER RULES
 *
 *  Order matters: more specific tokens must precede WORD.
 * ================================================================ */

BEGIN_DELIMITER : '=== BEGIN DISCHARGE SUMMARY ===' ;
END_DELIMITER : '=== END DISCHARGE SUMMARY ===' ;

/* De-identified PHI placeholder, e.g. [**2190-5-16**], [**Hospital1 69**],
   [**Last Name (STitle) 70**], [**Telephone/Fax (1) 170**] ... */
DEID
    : '[**' (~[\]])* '**]'
    ;

/* Time like 08:16 or 14:13:30 */
TIME
    : DIGIT DIGIT ':' DIGIT DIGIT (':' DIGIT DIGIT)? ('AM'|'PM'|'am'|'pm')?
    ;

/* Age phrase used by HPI intro: "51 year-old" / "73 yo" handled at
   the parser level using the AGE token below for the numeric piece. */
AGE
    : DIGIT DIGIT?      /* matched in context (parser); falls through to NUMBER
    otherwise */
    ;

PERCENT : '%' ;

NUMBER
    : DIGIT+ ('.' DIGIT+)?
    | DIGIT+ ',' DIGIT DIGIT DIGIT  /* e.g. 201,000 */
    ;

WORD
    : LETTER (LETTER | DIGIT | '\'' | '.')*
```

```
    ;

fragment LETTER : [a-zA-Z] ;
fragment DIGIT : [0-9] ;

WS  : [ \t\r\n]+ ;
ANY_CHAR : . ;
```

References

[1] J. F. Golob Jr, J. J. Como, and J. A. Claridge. "The painful truth: The documentation burden of a trauma surgeon". en. In: *J. Trauma Acute Care Surg.* 80.5 (May 2016), 742–5, discussion 745–7. DOI: 10.1097/TA.0000000000000986.

[2] B. G. Arndt, J. W. Beasley, M. D. Watkinson, J. L. Temte, W.-J. Tuan, C. A. Sinsky, and V. J. Gilchrist. "Tethered to the EHR: Primary care physician workload assessment using EHR event log data and time-motion observations". en. In: *Ann. Fam. Med.* 15.5 (Sept. 2017), pp. 419–426. DOI: 10.1370/afm.2121.

[3] O. Kleinig, L. Hains, A. Murugappa, S. Gluck, J. Marks, T. Gilbert, and S. Bacchi. "Time for a rethink? Discharge summary completion is often delayed and associated with increased readmission". In: *Internal Medicine Journal* 55.9 (2025), pp. 1575–1577. DOI: 10.1111/imj.70150.

[4] D. J. Morgan, W. Branch-Elliman, and K. E. Goodman. "Time to Study Implementation of AI-Generated Discharge Summaries". In: *JAMA Internal Medicine* 185.7 (July 2025), pp. 825–826. ISSN: 2168-6106. DOI: 10.1001/jamainternmed.2025.0827.

[5] K. Singhal et al. "Large language models encode clinical knowledge". In: *Nature* 620.7972 (Aug. 2023), pp. 172–180. ISSN: 1476-4687. DOI: 10.1038/s41586-023-06291-2.

[6] D. Van Veen et al. "Adapted large language models can outperform medical experts in clinical text summarization". In: *Nature Medicine* 30.4 (Apr. 2024), pp. 1134–1142. ISSN: 1546-170X. DOI: 10.1038/s41591-024-02855-5. URL: https://doi.org/10.1038/s41591-024-02855-5.

[7] C. Y. K. Williams et al. "Physician- and Large Language Model–Generated Hospital Discharge Summaries". In: *JAMA Internal Medicine* 185.7 (July 2025), pp. 818–825. ISSN: 2168-6106. DOI: 10.1001/jamainternmed.2025.0821.

[8] E. Asgari, N. Montaña-Brown, M. Dubois, S. Khalil, J. Balloch, J. A. Yeung, and D. Pimenta. "A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation". In: *npj Digital Medicine* 8.1 (May 2025), p. 274. ISSN: 2398-6352. DOI: 10.1038/s41746-025-01670-7.

[9] C. Y. K. Williams, J. Bains, T. Tang, K. Patel, A. N. Lucas, F. Chen, B. Y. Miao, A. J. Butte, and A. E. Kornblith. "Evaluating large language models for drafting emergency department encounter summaries". In: *PLOS Digital Health* 4.6 (June 2025), pp. 1–14. DOI: 10.1371/journal.pdig.0000899. URL: https://doi.org/10.1371/journal.pdig.0000899.

[10] M. Ganzinger, N. Kunz, P. Fuchs, C. K. Lyu, M. Loos, M. Dugas, and T. M. Pausch. "Automated generation of discharge summaries: leveraging large language models with clinical data". In: *Scientific Reports* 15.1 (May 2025), p. 16466. ISSN: 2045-2322. DOI: 10.1038/s41598-025-01618-7.

[11] S. Ellershaw et al. "Automated Generation of Hospital Discharge Summaries Using Clinical Guidelines and Large Language Models". In: *AAAI 2024 Spring Symposium on Clinical Foundation Models*. 2024. URL: https://openreview.net/forum?id=1kDJJPppRG.

[12] F. Grolleau et al. "Physician-Reported Safety Outcomes of AI-Generated Hospital Course Summaries". In: *JAMA Network Open* 9.5 (2026), e2616556. DOI: 10.1001/jamanetworkopen.2026.16556. URL: https://doi.org/10.1001/jamanetworkopen.2026.16556.

[13] F. L. Piya and R. Beheshti. *AgenticSum: An Agentic Inference-Time Framework for Faithful Clinical Text Summarization*. 2026. arXiv: 2602.20040 [cs.CL]. URL: https://arxiv.org/abs/2602.20040.

[14] A. Johnson, L. Bulgarelli, T. Pollard, B. Gow, B. Moody, S. Horng, L. A. Celi, and R. Mark. *MIMIC-IV (version 3.1)*. PhysioNet. RRID:SCR_007345. 2024. DOI: 10.13026/kpb9-mt58.

[15] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals". en. In: *Circulation* 101.23 (June 2000), E215–20.

[16] A. E. W. Johnson et al. "MIMIC-IV, a freely accessible electronic health record dataset". In: *Scientific Data* 10.1 (Jan. 2023), p. 1. ISSN: 2052-4463. DOI: 10.1038/s41597-022-01899-x.

[17] A. Johnson, T. Pollard, and R. Mark. "MIMIC-III Clinical Database". In: *PhysioNet* (Sept. 2016). Version 1.4. DOI: 10.13026/C2XW26. URL: https://doi.org/10.13026/C2XW26.

[18] A. E. Johnson, T. J. Pollard, L. Shen, L.-w. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Anthony Celi, and R. G. Mark. "MIMIC-III, a freely accessible critical care database". In: *Scientific Data* 3.1 (May 2016), p. 160035. ISSN: 2052-4463. DOI: 10.1038/sdata.2016.35. URL: https://doi.org/10.1038/sdata.2016.35.

[19] Health Level Seven International. *Clinical Quality Language (CQL) Specification*. https://cql.hl7.org/01-introduction.html. Accessed: May 2026. 2025.

[20] M. Samwald, K. Fehre, J. de Bruin, and K.-P. Adlassnig. "The Arden Syntax standard for clinical decision support: Experiences and directions". In: *Journal of Biomedical Informatics* 45.4 (2012). Translating Standards into Practice: Experiences and Lessons Learned in Biomedicine and Health Care, pp. 711–718. ISSN: 1532-0464. DOI: 10.1016/j.jbi.2012.02.001.

[21] Health Level Seven International. *FHIR Overview – Developers*. https://www.hl7.org/fhir/overview-dev.html. Accessed: May 2026. 2023.

[22] R. H. Dolin, L. Alschuler, C. Beebe, P. V. Biron, S. L. Boyer, D. Essin, E. Kimber, T. Lincoln, and J. E. Mattison. "The HL7 Clinical Document Architecture". en. In: *Journal of the American Medical Informatics Association* 8.6 (Nov. 2001), pp. 552–569. DOI: 10.1136/jamia.2001.0080552.

[23] Health Level Seven International. *FHIR Example of a discharge summary*. https://www.hl7.org/fhir/document-example-dischargesummary.json.html. Accessed: May 2026. 2023.

[24] Health Level Seven International. *HL7 Clinical Document Architecture (CDA)*. https://hl7.org/cda/. Accessed: May 2026. 2025.

[25] J. Ma and M. Agrawala. "MoVer: Motion Verification for Motion Graphics Animations". In: *ACM Transactions on Graphics* 44.4 (2025), pp. 1–17. DOI: 10.1145/3731209.

[26] B. Li, R. Wu, A. Solar-Lezama, L. Shi, C. Zheng, B. Bickel, and W. Matusik. "VL-Material: Procedural Material Generation with Large Vision-Language Models". In: *The Thirteenth International Conference on Learning Representations*. 2025. URL: https://openreview.net/forum?id=wHebuIb6IH.

[27] L. Makatura, B. Jones, S. Bian, and W. Matusik. *MetaGen: A DSL, Database, and Benchmark for VLM-Assisted Metamaterial Generation*. 2025. arXiv: 2508.17568 [cs.CV]. URL: https://arxiv.org/abs/2508.17568.

[28] J. Li, W. Ma, X. Li, Y. Lou, G. Zhou, and X. Zhou. "CAD-Llama: Leveraging Large Language Models for Computer-Aided Design Parametric 3D Model Generation". In: *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2025, pp. 18563–18573. DOI: 10.1109/CVPR52734.2025.01730.

[29] S. Agarwal, L. Ahmad, J. Ai, S. Altman, et al. *gpt-oss-120b & gpt-oss-20b Model Card*. 2025. DOI: 10.48550/arXiv.2508.10925. arXiv: 2508.10925 [cs.CL]. URL: https://arxiv.org/abs/2508.10925.

[30] G. Gerganov and contributors. *llama.cpp: LLM inference in C/C++*. https://github.com/ggml-org/llama.cpp. Accessed: May 2026. 2023.

[31] O. Khattab et al. "DSPy: Compiling Declarative Language Model Calls into State-of-the-Art Pipelines". In: *The Twelfth International Conference on Learning Representations*. 2024. DOI: 10.48550/arXiv.2508.10925. URL: https://openreview.net/forum?id=sY5N0zY5Od.

[32] L. A. Agrawal et al. *GEPA: Reflective Prompt Evolution Can Outperform Reinforcement Learning*. 2025. arXiv: 2507.19457 [cs.CL]. URL: https://arxiv.org/abs/2507.19457.

[33] G. LoDuca. *Introducing Parley: MIT's Generative AI Platform*. https://orgchart.mit.edu/letters/introducing-parley-mits-generative-ai-platform. Accessed: May 2026. 2026.

[34] PhysioNet. *Responsible Use of MIMIC Data with Online Services Like GPT*. https://physionet.org/news/post/gpt-responsible-use/. Accessed: May 2026. Apr. 2023.