The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200525

A Comparative Study of the Arden Syntax and GDL Clinical Knowledge Representation Languages

Jeroen S. DE BRUIN [a,1], Rong CHEN [b,c],
Andrea RAPPELSBERGER [d] and Klaus-Peter ADLASSNIG [d,e]
[a] *Institute eHealth, FH Joanneum University of Applied Sciences GmbH, Graz, Austria*
[b] *Cambio Healthcare Systems, Stockholm, Sweden*
[c] *Health Informatics Centre, LIME, Karolinska Institute, Solna, Sweden*
[d] *Section for Artificial Intelligence and Decision Support, Medical University of Vienna, Vienna, Austria*
[e] *Medexter Healthcare GmbH, Vienna, Austria*

Abstract. The expressiveness of a medical knowledge representation language has significant impact on the effectiveness of a knowledge-based clinical decision support system. We assess the expressiveness of two such languages, Arden Syntax and the Guideline Definition Language. Using data extracted from both languages' specifications, we quantify expressiveness by means of language syntax and the number of supported operators. Preliminary results show that Arden Syntax is a more dynamic standard, having better readability and a higher number and more diverse operators than GDL. In contrast, GDL is a more rigid language that utilizes an underlying data model specification in the openEHR framework.

Keywords. Decision Support Systems, Clinical; Arden Syntax; Guideline Definition Language

1. Introduction

With knowledge-based clinical decision support (CDS) approaches, knowledge is implemented through the extraction and translation of medical knowledge sources using a knowledge representation language. The expressiveness of the knowledge representation language, i.e., the variety and quantity of logics, ideas and rules that it can express, thereby has significant impact on the effectiveness of CDS systems.

The aim of this paper is to present preliminary results on a study to assess and compare the expressive power of knowledge representation languages for knowledge-based CDS systems. At present, two such languages are commonly used, the Arden Syntax for Medical Logic Modules (Arden Syntax for short) [1] and the Guideline Definition Language (GDL) [2] of the openEHR framework.

[1] Corresponding Author, Drs. Dr. Jeroen S. de Bruin, Institute eHealth, FH Joanneum University of Applied Sciences GmbH, Eckertstrasse 30i, 8020 Graz, Austria; E-Mail: jeroen.debruin@fh-joanneum.at.

2. Methods

We performed a comparative study of the expressiveness of knowledge representation languages for CDS systems, Arden Syntax and GDL. Both languages have an open specification, are still further developed in workgroups, have reference implementations available and have been used for the implementation and publication of CDS systems [3–6].

2.1. The Arden Syntax

Arden Syntax is a standard for the representation and processing of clinical and scientific knowledge managed by Health Level Seven (HL7) International, a non-profit ANSI-accredited standards development organization for health data interoperability [1].

Meant to be a standard for use in the medical domain, the Arden Syntax offers data types and constructs tailored to the need of medical documentation and decision making, e.g., time series and duration. Moreover, Arden Syntax resembles natural language, which improves its readability and understandability.

In Arden Syntax, each knowledge base is partitioned into medical logic modules (MLMs), which are independent units that contain enough knowledge to compute at least one medical decision. Each MLM is divided into four categories that contain maintenance information, links to other sources of knowledge, and logic to make a health decision, and optional localization for return messages.

2.2. openEHR and the Guideline Definition Language

openEHR is an open standard for the storage, retrieval, management and exchange of health data in electronic health records (EHRs) [7]. Within the openEHR program, specifications of formal models and languages for the definition of the openEHR technical platform are generated, clinical models are created for re-usable clinical content, and software is created that implement open source tools and EHR information content.

GDL [2] is a formal language for expressing decision support logic in openEHR. It is designed to be agnostic to natural languages and references terminologies by leveraging openEHR foundation types and formalisms. The standard was developed by Cambio Healthcare Systems and was later adopted by the openEHR Foundation.

Knowledge-based CDS applications created with GDL have a specific object model that consists of a resource description, terminology definitions and bindings, and the rules that define its functionality. Only the resource description and the terminology definition parts contain labels in natural language. The terminology definition part of the guideline acts as an "interface" between natural language and the purely coded CDS rules.

2.3. Study Flow & Outcome Measures

We first established comparison criteria for selected knowledge representation languages. Then, we collected general information on these languages. Information on Arden Syntax was collected from the latest specification and implementation guide [1,8]; information on GDL was collected from the openEHR GDL specification site [2] and the project website [9]. Finally, we extracted information on selected comparison criteria from these information sources.

As outcome measures, we discuss language syntax and the number of operators supported by the language. Although to the authors' knowledge a formal comparison framework does not exist, discussing these aspects will give a quantitative indication of language expressiveness.

3. Results

A comparison of syntax attributes between Arden Syntax and GDL is shown in Table 1. Arden Syntax is designed with readability in mind and its syntax resembles that of modern object-oriented programming languages. In contrast, GDL is designed to be language-agnostic, and its syntax resembles that of markup languages such as XML.

Table 1. Syntax attributes for Arden Syntax and the Guideline Definition Language.

Attribute	Arden Syntax	Guideline Definition Language
Readability	Based on English natural language, readable for technical and nontechnical users.	Language-agnostic, hardly involves any natural language.
Syntax	Resembling functional and object-oriented programming languages.	Resembling XML markup languages.
Keywords	Some reserved keywords, variable and parameter names can be chosen freely otherwise.	Variables and parameters bound to data points with an archetype.

With respect to the types and number of available operators (Table 2), Arden Syntax shows more diversity that GDL; the Arden Syntax supports both Boolean and fuzzy logics and supports 176 operators. In comparison, GDL defines a total of 28 operators in its specification.

Table 2. Operators for the Arden Syntax and the Guideline Definition Language specified by operator categories.

Operator category	Arden Syntax	Guideline Definition Language
Logical Operators	3	3
Comparison Operators*	44	6
Arithmetic Operators	7	6
Function Operators**	17	11
List operators, filters, string operations, temporal operators, aggregations, transformations, object operators, fuzzy operators, type conversions	105	
Terminological reasoning operators		2
Total	**176**	**28**

* Simple comparisons, Is comparisons, and Occur comparisons for Arden Syntax

** Numeric and time functions for Arden Syntax

4. Discussion

We performed a study on language expressive power between Arden Syntax and GDL by comparing language syntax and the number of operators supported by the language. While the choice for a CDS language is not solely determined by these factors, and the purpose and context of CDS also play a significant role, this study provides a quantitative measure to support language choice.

Arden Syntax and GDL have two different philosophies, which clearly shows in their respective syntaxes. Arden Syntax is designed with readability in mind, whereas GDL is designed to be language-agnostic. This makes it easier for programmers to learn Arden Syntax, and for clinicians to verify rules expressed in MLMs. In comparison, GDL implementations are more difficult to read and understand.

The difference in available operator types and functions in the two CDS languages mirrors the more flexible approach of Arden Syntax versus the strictly standardized approach of GDL. However, it should be noted that the effects of some Arden Syntax operators can be achieved by combining GDL operators with appropriate data types.

5. Conclusions

Both Arden Syntax and GDL are languages for the creation of CDS applications. Arden Syntax is more flexible and offers more possibilities of creating complex algorithms. GDL is a more rigid standard that guarantees its function when implemented into a system that is based on the openEHR specifications.

References

[1] Health Level Seven International, HL7 Standards Product Brief – Arden Syntax v2.10 (Health Level Seven Arden Syntax for Medical Logic Systems, Version 2.10), https://www.hl7.org/implement/standards/product_brief.cfm?product_id=372, last access: 29.01.2020.

[2] The openEHR Foundation, Guideline Definition Language (GDL), https://specifications.openehr.org/releases/CDS/latest/GDL.html, last access: 13.03.2020.

[3] Flores B, Design and Development of a Decision Support System for Screening of Lynch Syndrome Using openEHR, Master Thesis, Karolinska Institutet, Stockholm, 2015.

[4] Anani N, Chen R, Moreira TP, Koch S. Representative checking of compliance with practice guidelines for acute stroke care: A novel experiment using openEHR's Guideline Definition Language, BMC Medical Informatics and Decision Making 14 (2014), 39. https://doi.org/10.1186/1472-6947-14-39.

[5] de Bruin JS, Adlassnig KP, Blacky A, Mandl H, Fehre K, Koller W, Effectiveness of an automated surveillance system for intensive care unit-acquired infections, Journal of the American Medical Informatics Association 20(2) (2013), 369–372. https://doi.org/10.1136/amiajnl-2012-000898.

[6] Schuh C, de Bruin JS, Seeling W, Clinical decision support systems at the Vienna General Hospital using Arden Syntax: Design, implementation, and integration. Artificial Intelligence in Medicine 92 (2018), 24–33. https://doi.org/10.1016/j.artmed.2015.11.002.

[7] The openEHR Foundation, What is openEHR?, https://specifications.openehr.org/releases/CDS/latest/GDL.html, last access: 13.3.2020.

[8] Arden Syntax Work Group, HL7 Arden Syntax: Implementation Guide Release 2, HL7 Informative Ballot, April 2017, https://wiki.hl7.org/index.php?title=Arden_Syntax_Implementation_Guide_Release_2, last access: 29.3.2020.

[9] Medical Informatics Group at Cambio Healthcare Systems, Guideline Definition Language (GDL), https://gdl-lang.org/, last access: 23.3.2020.