論 文
55D-6-4

효율적인 HL7 메시지 파싱 알고리즘

An Efficient Algorithm for HL7 Message Parsing

트란 뚱* · 金 和 仙† · 金 亨 會** · 趙 勳***
(Tung Tran · Hwa-Sun Kim · Hyung-Hoi Kim · Hune Cho)

Abstract – An upgraded algorithm that improves the performance of existing interfacing software for parsing HL7 messages is introduced. It incorporates stack operations on objects to guarantee segment order while parsing messages. This object-oriented design greatly facilitates the complicated process of validating, parsing, and creating HL7 messages in the clinical setting. The new interface engine can manage all HL7 messages corresponding to admission and registration, discharge and transfer, laboratory results, clinical images, and clinical reports. The international version of this engine, currently under development, will be tested in Asian countries using standard character code such as Unicode (ISO 10646).

Key Words : HL7, Interface Engine, Parsing Algorithm

1. INTRODUCTION

Hospitals, doctors, healthcare professionals, and institutions regularly exchange information, and the HL7 standard was written to facilitate this communication [1]. HL7 is based on approximately 300 real-world trigger events, and its specifications are defined in terms of different message types, each with a distinct purpose [2]. Message types may describe registration, results and observations, orders, and other information. The HL7 interface engine is considered a critical component of the IT infrastructure in a hospital setting, because it coordinates the data elements of every system [3, 4]. For example, admission, discharge, and transfer data are sent from the registration system to the patient care information, clinical information, medical records, and billing systems. Also it interfaces clinical information results with patient care results reporting. Unfortunately, however, most commercial HL7 interfaces are expensive, anywhere from hundreds to thousands of dollars per individual license, and are limited to the vendor's own product. For these reasons, Kyungpook National University School of Medicine, Department of Medical Informatics, developed an HL7 interface engine using a streaming algorithm. Called Clinical Primary Exchanger version 1.0 (CPX 1.0), the interface was rudimentary, but sufficient for exchanging discharge summaries between two hospitals [5]. In this paper, we introduce an improved algorithm that incorporates an object-oriented design and stack operations to enhance the efficiency of the CPX 1.0 interface.

† 교신저자, 正 會 員 : 慶北大學校 醫療情報學科 博士課程
E-mail : pulala@paran.com
* 正 會 員 : 慶北大學校 醫療情報學科 博士課程
** 正 會 員 : 釜山大學校 醫科大學 診斷檢査醫學科 敎授
*** 正 會 員 : 慶北大學校 醫科大學 醫療情報學科 敎授
接受日字 : 2006年 3月 27日
最終完了 : 2006年 5月 12日

2. METHODS

2.1 HL7 message

HL7 messages are constructed as a list of segments with data fields that are either HL7 primitive data types or HL7 composite data types [6, 7]. Each field contains data elements known as components, which may be further divided into sub-components. Fields may repeat, and the maximum number of repetitions may also be repeated. Data types may be simple, such as date (DT), telephone number (TN), and identifier (ID), or compound, such as addresses, phone numbers, and patient names. Delimiter characters separate the segments (Table 1).

There are more than 100 messages currently defined, including ADT messages, laboratory test results, and some query messages. The message header segment (MSH) provides necessary information regarding each HL7 message, which is unique based on its segment order.

Table 1 Delimiter characters

Segment terminator	<cr> 0x0D
Field Separator	|
Component Separator	^
Sub-Component Separator	&
Repetition Separator	~
Escape Separator	\

2.2 Object-oriented programing

Object-oriented programming (OOP) is a new methodology that models real-world concepts as objects with both source and executable code [8]. Objects are comprised of two parts: data, and the operations that can be performed on these data. So each HL7 message is based on a real-world event, and contains data and operations for that event. Type Interface is a well-defined HL7 data type [9] implemented by either primitive or composite interfaces (Fig. 1). We have also designed Primitive Interface, Composite Interface, Segment Interface, and Group Segment Interface, which contain primitive data types, composite data types, segments, and segment groups, respectively, within their attributes, as defined by HL7 Standard [10].

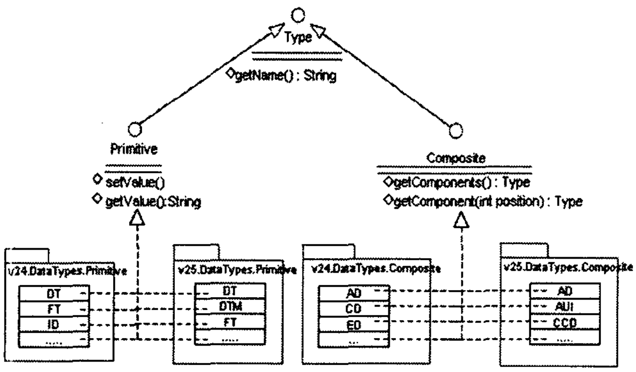

Fig. 1 Encapsulation and iheritance using HL7 data types

2.3 Parse algorithm

Parsing is the process of determining whether a string of tokens can be generated by a grammar. The parser should report any errors in an intelligible fashion, and recover from common errors so that it can continue processing the remainder of its input. In the proposed HL7 interface engine, the parse checks message structure and validate data type (Fig. 2).

For each incoming message, the system checks the MSH, and if it is valid, instantly initiates the parsing process, which proceeds as follows. The message object is loaded by retrieving the message type and event from MSH-9, and then the object is checked to see if it exists

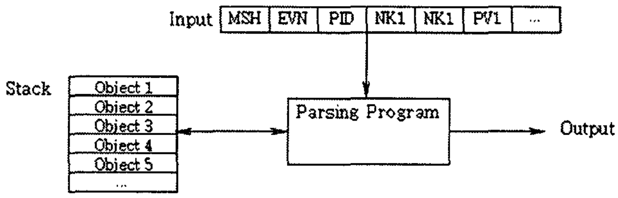

Fig. 2 Model of parse

on the system. If so, the system loads the message object, retrieves its segments, and stores them in a stack, which orders the segments. If, on the other hand, the message object is not validated on the system and thus not loaded, the system will record an internal error and return a rejection message. Parse processing consists of message, segment, and field parsing (Fig. 3).

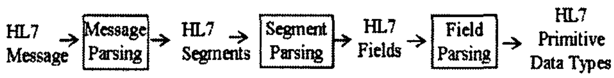

Fig. 3 parsing of an HL7 message

An example of message parsing follows. The segment and field parsing algorithms are similar (Fig. 4).

```
Procedure MessageParsing
Loop
  pop object from Stack
  read segment from input array
  if segment.name matches object.segment.name then
    call parseSegment procedure;
    if object.segment can be repeated then
      loop
        read next incomingMes.segment
        if next segment matches object.segment then
              create and add new object segment to message object
      until segment doesn't match object.segment or input is empty
    end if
  end if
  else
    if object.segment is required then
      add Error message to ErrorList
    end if
  end else
until Stack is empty
end procedure
```

Fig. 4 Segment and field parsing algorithms

3. RESULTS

The main module of this interface engine includes message validation, browsing, generation, sending, acknowledgment, and parsing. In a test run, a discharge summary of XML-encoded clinical document architecture (CDA) [11-14] confirmed the possibility of exchange in the simulated environment (Fig. 5).

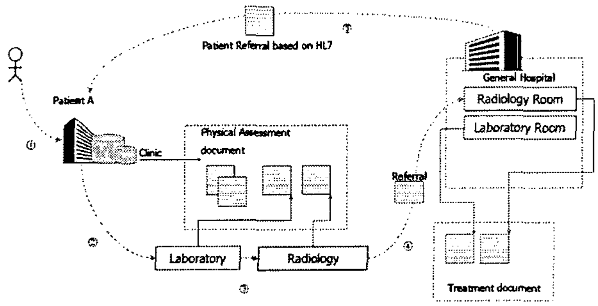

Fig. 5 Transaction flows in a simulated environment: *a)* Patient A presents to the clinic with dizziness and epigastric pain; *b)* Physician sees the patient, *c)* performs a CBC, LFT, X-ray, and GFS, *d)* diagnoses internal bleeding due to a gastric ulcer, and refers the patient to the general hospital; *e)* At the general hospital, the emergency room doctor retrieves and reviews the relevant reports; *f)* After intervention, Patient A returns to the physician at the clinic for a follow-up visit

Performance evaluation: Table 2 shows the parsing time of CPX 2.0 and CPX 1.0. Table 3 compares the CPX 1.0, CPX 2.0, and jEngine systems. In the past, CPX 1.0 is tested within transferring discharge summary. The message contains laboratory test results, patient demographic information. However, CPX 2.0 is tested within ADT messages, discharge summary, and DICOM images.

Table 2 Parsing Time of CPX 1.0 and CPX 2.0

Size of the Message(KB)	CPX 1.0(ms)	CPX 2.0(ms)
94.4	560	78
115.9	750	170
185.1	1800	192
456.6	4150	310
537.4	6500	351

Table 3 Comparison of *j*Engine, CPX 1.0 and CPX 2.0

	*j*Engine	CPX 1.0	CPX 2.0
Validate MSH segment	Yes	No	Yes
Check required segment	No	No	Yes
Check field's length	No	No	Yes
Validate message syntax	No	Yes	Yes
Check required field	No	Yes	Yes
Validate data type	Yes	Yes	Yes
Exchage information among interface	Yes	No	Yes
HL7 messages	All	ADT_A03	ADT, OBX
Generate ACK message within Error Segment	No	Yes	Yes

4. CONCLUSION

4.1 Message parsing

A received message can normally be parsed into message objects that are already defined in HL7 messaging, making it simple to check the message structure, field length, required segments, and to validate the datatype value (length and data type format). In addition, this system can generate an error message when one occurs during the parsing process (Fig. 6).

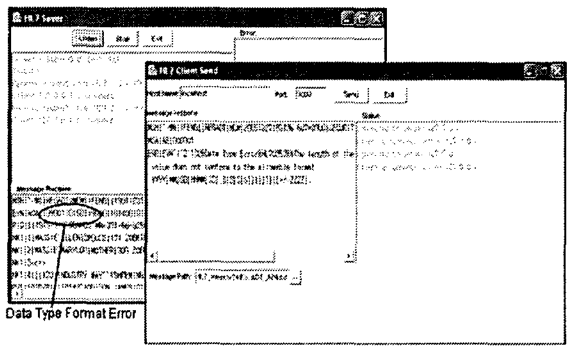

Fig. 6 An example of an error message

4.2 Text and image data transfer

CPX's message browser can display HL7 message structure in tree form. Because this structured view is more easily understood than the HL7 message itself, it is important for HL7 interface engine programming and education training that the message structure established by parsing is displayed. The proposed system supports this tree-browsing method. How can reporting needs be met in terms of continuity of information? Critical factors include ease of access and availability of relevant information. Information relevance is context-dependent, determined by the study context of the radiological procedure (current and prior imaging studies and reports) and the patient context, where procedure and indication-based access to clinical information is required. While this calls for the integration of image and report information as well as relevant clinical information other than imaging, current limitations primarily stem from the use of hardcopy images and paper reports and transcribed narrative text reports with no imaging interpretation results included or referenced.

5. DISCUSSION

The HL7 message standard is remarkably powerful for such a limited design. Furthermore, this simple parser could easily be expanded to both encode and parse HL7 data. The fact that plain text can hold such large amounts of data clearly shows that medical information sharing need not be hindered by technical limitations. The technology needed to solve many of the information problems in the medical field is already available. What is needed is for more healthcare professionals, those who conduct their daily business in medicine, to add their input to the further development of applications. Because they know best what the needs and special considerations of their work are, it is imperative that their ideas be considered in the development cycle. Hopefully, the proposed program described herein contributes to this effort, even if only to spark the interest of those in the HL7, medical messaging, and health information fields.

The proposed system improves the parsing method of HL7 messages, utilizing an object-oriented design and stacks. This new HL7 interface engine enhances network efficiency during HL7 message transfer, whether clinical images or CDA. The engine also successfully generated, sent, received, parsed, and detected images automatically included in HL7 messages (versions 2.4 and 2.5). The international version of this engine is currently under development [15].

ACKNOWLEDGMENTS

This work was supported by grant No. (R-2006-000-10009-0) from the Basic Research Program of the Korea Science & Engineering Foundation and No. (718) from the Dong-il Research Program of the Dong-il Culture Foundation.

REFERENCES

[1] J. Lira, "High standards. The formidable task ofadopting HL7 standards is easier with the right tools," Healthc Inform, vol. 17, pp. 131-2, 2000.

[2] R. Hussein, U. Engelmann, A. Schroeter, and H. P. Meinzer, "Implementing a full-feature PACS solution in accordance with the IHE technical framework: the CHILI approach," Acad Radiol, vol. 11, pp. 439-47, 2004.

[3] G. Schadow, U. Fohring, and T. Tolxdorff, "Implementing HL7:from the standard's specification to production application," Methods Inf Med, vol. 37, pp. 119-23, 1998.

[4] P. Mildenberger, B. Wein, H. P. Bursig, and M. Eichelberg, "[Current developments in DICOM and IHE]," Radiologe, vol. 45, pp. 682-9, 2005.

[5] K. S. Um, Y. S. Kwak, H. Cho, and I. K. Kim, "Development of an HL7 interface engine, based on tree structure and streaming algorithm, for large-size messages which include image data," Comput Methods Programs Biomed, vol. 80, pp. 126-40, 2005.

[6] R. A. Jenders and R. M. Sailors, "Convergence on a standard for representing clinical guidelines: work in health level seven," Medinfo, vol. 11, pp. 130-4, 2004.

[7] J. Quinn, "An HL7 (Health Level Seven) overview," J Ahima, vol. 70, pp. 32-4; quiz 35-6, 1999.

[8] A. Beveridge, "An object-oriented programming system for the integration of internet-based bioinformatics resources," Appl Bioinformatics, vol. 5, pp. 29-39, 2006.

[9] M. Koncar, "[HL7 standard--features, principles, and methodology]," Acta Med Croatica, vol. 59, pp. 273-6, 2005.

[10] E. W. Huang, D. W. Wang, and D. M. Liou, "Development of a deterministic XML schema by resolving structure ambiguity of HL7 messages," Comput Methods Programs Biomed, vol. 80, pp. 1-15, 2005.

[11] R. H. Dolin, L. Alschuler, S. Boyer, C. Beebe, F. M. Behlen, P. V. Biron, and A. Shabo Shvo, "HL7 Clinical Document Architecture, Release 2," J Am Med Inform Assoc, vol. 13, pp. 30-9, 2006.

[12] J. M. Ferranti, R. C. Musser, K. Kawamoto, and W. E. Hammond, "The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis," J Am Med Inform Assoc, 2006.

[13] H. B. Bludau, A. Wolff, and A. J. Hochlehnert, "Presenting XML-based medical discharge letters according to CDA," Methods Inf Med, vol. 42, pp. 552-6, 2003.

[14] C. Loef and R. Truyen, "Evidence and diagnostic reporting in the IHE context," Acad Radiol, vol. 12, pp. 620-5, 2005.

[15] E. W. Huang, S. H. Hsiao, and D. M. Liou, "Design and implementation of a web-based HL7 message generation and validation system," Int J Med Inform, vol. 70, pp. 49-58, 2003.

저 자 소 개

트란 둥(Tung Tran)

Tung Tran, MR is received M.S degreein Medical Informatics from Kyungpook Nation University. PhD. student at Department of Medical Informatics, Kyungpook National University, Daegu, Republic of Korea. His major interests is design and implementation of "parsing algorithm", and HL7 interface engine. Currently he has been working on "design and implementation of EMR system for small practice."Interested readers may contact the author via either ttadm20031@yahoo.com or Department of Medical Informatics at Kyungpook National University Shool of Medicine, Jung-gu Dongin-dong, 2-101, daegu Korea, 700-422

김 화 선(金 和 仙)

Ms Hwa Sun Kim, MS is a researcher and PH.D. student at Department of Medical Informatics, Kyungpook National University, Daegu, Republic of Korea. She is s RN and has the M.S. degree in Computer Science and Engineering. Prior to joining the current position, she has worked as a Medical Underwriter at Claims & Investigation Department of Allianz Life Insurance Co. Ltd, a Computer Systems Applications and Development Engineer and Team leader of Nursing Information System Design for Hospital Information System and a Chief Nurse Mary Knoll Hospital, Busan, Korea. Her major interests and past accomplishments are design and implementation of XML based systems. Currently she has been working on "Architectural design of new paradigm hospital information system based on web service."

김 형 회(金 亨 會)

Hyung Hoi Kim, M.D.,PhD is a professor of the Unit of Biomedical Informatics and Department of Laboratory Medicine at Pusan National Univ. School of Medicine. He was born and graduated school of medicie, Pusan National University, Korea. he teaches varieties of laboratory medicine and medical information. His major research interests are hospital information systems, Laboratory information, web-based telemedicine and many other health related issues. Interested readers may contact the author, via either hhkim@pusan.ac.kr or Department of laboratiory medicine at Pusan National University School of Medicine, Seo-gu Ami-dong , Busan Korea, 602-739.

조 훈(趙 勳)

Hune Cho, PhD is a professor of the Department of Medical Informatics at Kyungpook National Univ. School of Medicine. He was born and educated in Korea and further research was continued in USA. On his return, he teaches varieties of medical informatics courses and initiates research on information systems of Kyungpook University Hospital. In addition, he is currently involved in international standardization of medical information for ISO/TC215. His major resarch interests are hospital information systems, medical imaging, computerzied patient record, web-based telemedicine and many other health related issues. Interested readers may contact the author, via either hunecho@knu.ac.kr or Department of Medical Informatics at Kyungpook National University School of Medicine, Jung-gu Dongin-dong 2-101, Taegu Korea, 700-422.