

Toward the Discovery and Extraction of Money Laundering Evidence from Arbitrary Data Formats using Combinatory Reductions

Alonza Mumford, Duminda Wijesekera

George Mason University

amumford@gmu.edu, dwijesek@gmu.edu

November 19, 2014

- 1 Problem, Background & Architecture
- 2 Motivating Example
- 3 Descriptive Data Modeling with DFDL
- 4 Ontological Modeling with XLink
- 5 Compiler Specification in CRSX
- 6 Future Work

Problem Statement

- The "wild, wild west analogy" is applicable to the variety of non-standardized, structured and unstructured data that exist in public and private domains
- This phenomena disrupts the opportunity to "get to" this heterogeneous raw data in a generic manner, represent the connections between this data that is distributed, and make new discovers based partially on those connections.
- The manner in which we express these connections must contextualize specific data with a conceptual understanding of a problem, investigation or inquiry without having to change or alter the data itself

Background (Challenges)

Challenge 1

To provide a mechanism that can be used to describe and access any number of data formats.

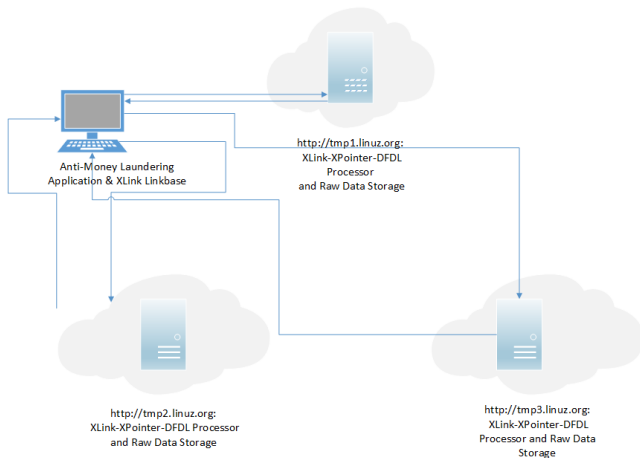
Challenge 2

To provide a lightweight, metadata-based discovery and extraction of arbitrary data fragments from raw data stores not co-located with each other. The intent is to avoid some of the system development and maintenance costs associated with major data conversion, and database storage and indexing.

Challenge 3

To provide a mechanism that incorporates ontological engineering for describing the meaning of data fragments (i.e., parts of data) within the context of a particular domain of understanding, inquiry or investigation.

Architecture (XLink-DFDL Processing)



Motivating Example: Analysis of a Money Laundering Scheme

Step 1: Placement Stage

A number of policies entered into by the same insurer (i.e. a person or company that underwrites an insurance risk) for small amounts and then canceled early at the same time [1, 2].

Step 2: Layering Stage

Requests for return premiums in currencies different to the original premium or requests for return premiums to an account different from the original account [1, 2].

Step 3: Integration Stage

Return premium being credited in different currency or different account, which represents launder money [1, 2].

Example (part 1): Inspection of a Policy Account Record

1 (a.1) Input "policy account record" data:

2

3 PLCYACC/741032-1071//

4 DATE/2013-09-28//

5 PLCYHLD/Allegier , Cox & Associates , Inc.//

6 INSUR/ALI Corp.//

7 PAYER/Grupo Palermo S.A.//

8 PAYCUR/Peso (ARG)//

9 PRMAMT/42004.98//

10

11 (a.2) DFDL generated XML model:

12

13 <policyAccountRecord>

14 <policyAccountIdentifier>741032-1071</
 policyAccountIdentifier>

15 <policyStartDate>2013-09-28</policyStartDate>

16 <policyHolder>Allegier , Cox & Associates Inc.</
 policyHolder>

17 <policyInsurer>ALI Corp.</policyInsurer>

18 <payerName>Grupo Palermo S.A.</payerName>

19 <payerCurrency>Peso (Argentine)</payerCurrency>

20 <premiumAmount>42004.98</premiumAmount>

21 </policyAccountRecord>

Example (part 2): Inspection of a Credit Request Record

1 (b.1) Input "credit request record" data:

2

3 [[[[[[[CRDREQ%]741032-1071%]

4 CRDATE%]2013-11-02%]

5 PAYEE%]Allegier , Cox & Associates , Inc.%]

6 PAYCUR%]USD%]

7 CRDAMT%]5000.00%]

8

9 (b.2) DFDL generated XML model:

10

11 <creditRequestRecord>

12 <creditRequestIdentifier>741032-1071</
 creditRequestIdentifier>

13 <creditRequestDate>2013-11-02</creditRequestDate>

14 <payeeName>Allegier , Cox & Associates , Inc.</payeeName>

15 <payeeCurrency>USD</payeeCurrency>

16 <creditAmount>5000.00</creditAmount>

17 </creditRequestRecord>

- Data Format Description Language [3] is not a data-format... it's a way to describe any data format (i.e., dense binary and text)
- Prescriptive standards (e.g. ASN.1, JSON) provides a structure for data as certain people think it should be in lieu of the structure of data as it actually is
- DFDL's syntax uses a subset of XML Schema with DFDL annotations to represent the physical properties of data
- As of January 2011, DFDL v1 is a "Recommendation" standard of the Open Grid Forum (OGF) designed to "enable efficient representation of data... essential to high performance data exchange in grid computing and... distributed computing."

Modeling of Logical Datatypes & Constraints

Logical Datatypes & Constraints	Context Free Grammar (CFG) example	Higher-Order Abstract Syntax (HOAS) example
Structures (xs:complexType)	<pre>2 <decl_xs> ::= "<" XS_COMPONENT <stmt>* ">/" XS_COMPONENT ">" 11 XS_COMPONENT ::= "xs:complexType"</pre>	XsComponent[ComplexType]
Atomic data values (xs:simpleType)	<pre>11 XS_COMPONENT ::= "xs:simpleType"</pre>	XsComponent[SimpleType]
Ordering (xs:sequence or xs:choice)	<pre>11 XS_COMPONENT ::= "xs:sequence" "xs:choice"</pre>	XsComponent[Sequence], XsComponent[Choice]
Occurrences (xs:minOccurs or xs:maxOccurs)	<pre>5 <stmt> ::= XS_ATTRIBUTE "=" <xs_attribute_value> 16 XS_ATTRIBUTE ::= "xs:minOccurs" "xs:maxOccurs" 8 <xs_attribute_value> ::= <xs_number></pre>	ComponentAttribute[MinOccurs], ComponentAttribute[MaxOccurs]

Modeling of Physical Representations

DFDL Physical Representation Properties	Context Free Grammar (CFG) example	Higher-Order Abstract Syntax (HOAS) example
Physical types (dfdl:representation)	<pre>5 <stmt> ::= DFDL_ATTRIBUTE "=" <dfdl_attribute_value> 18 DFDL_ATTRIBUTE ::= "representation" 7 <dfdl_attribute_value> ::= <dfdl_enum_number></pre>	FormatProperty[Representation]
Delimiters (dfdl:initiator, dfdl:separator, dfdl:terminator)	<pre>18 DFDL_ATTRIBUTE ::= "initiator" 7 <dfdl_attribute_value> ::= <dfdl_string_value> <reg_exp_value></pre>	FormatProperty[Initiator], FormatProperty[Separator], FormatProperty[Terminator]
Extraction of elements (dfdl:lengthKind)	<pre>18 DFDL_ATTRIBUTE ::= "lengthKind" 7 <dfdl_attribute_value> ::= <dfdl_enum_value></pre>	FormatProperty[LengthKind]
Points of uncertainty (dfdl:discriminator)	<pre>3 <dcl_dfdl> ::= "<" DFDL_ADMIN <stmt>* ">" <dcl_dfdl> "</" DFDL_ADMIN ">" 15 DFDL_ADMIN ::= "dfdl:discriminator" 18 DFDL_ATTRIBUTE ::= "occursCount"</pre>	DfdlValidation[Discriminator]
Detecting occurrences (dfdl:occursCount)	<pre>7 <dfdl_attribute_value> ::= <non_neg_int_value> <dfdl_exp_value></pre>	FormatProperty[OccursCount]

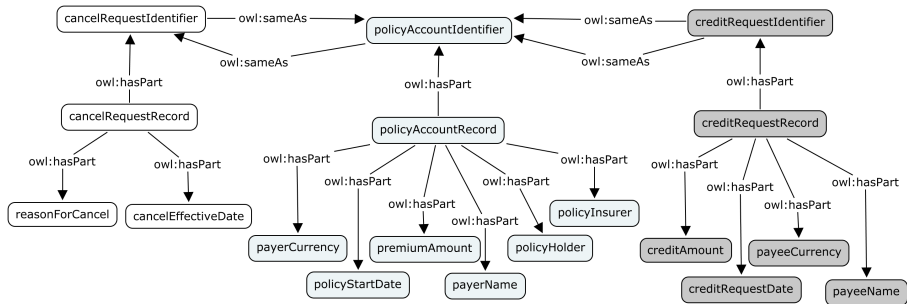
Example (part 3): Policy Account Record Schema

```
1  ...
2  <xs:element name="policyAccountRecord" minOccurs="0"
3    maxOccurs="unbounded" dfdl:lengthKind="implicit">
4    <xs:complexType>
5      <xs:sequence dfdl:sequenceKind="ordered">
6        <annotation>
7          <xs:appinfo source="http://www.ogf.org/dfdl
8            /v1.0">
9            <dfdl:element representation="text"
10              encoding="ascii" lengthKind="
11                delimited" sequenceKind="ordered"
12              initiator="/" separator="/"
13              separatorPosition="infix"
14              separatorPolicy="required"/>
15          </xs:appinfo>
16        </annotation>
17        <xs:element name="policyAccountIdentifier"
18          type="xs:string" dfdl:lengthKind="explicit"
19          dfdl:length="20"/>
20        <xs:element name="policyStartDate" type="
21          xs:string" dfdl:lengthKind="explicit"
22          dfdl:length="20"/>
23        <xs:element name="policyHolder" type="
24          xs:string" dfdl:lengthKind="explicit"
25          dfdl:length="20"/>
26        <xs:element name="policyInsurer" type="
27          xs:string" dfdl:lengthKind="implicit"/>
28        <xs:element name="payerName" type="xs:string"
29          dfdl:lengthKind="implicit"/>
30        <xs:element name="payerCurrency" type="
31          xs:string" dfdl:lengthKind="implicit"/>
32        <xs:element name="premiumAmount" type="
33          xs:string" dfdl:lengthKind="implicit"/>
34      </xs:sequence>
35    </xs:complexType>
36  </xs:element>
37  ...
```

Use XLink-based ontological engineering and analytical reasoning to define the concepts relevant to the money laundering domain [4, 5]:

- A *concept* definition conveys the name of an evidentiary fact and its value data type
- A *conceptual ontology* of the "anti-money laundering" domain is given to show the kinds of classes and properties used in the domain
- Classes* are identified by nodes and *properties* are identified by directed paths or arcs
- The *conceptual labels* associated with properties represent composition (i.e., hasPart) and equivalence (i.e., sameAs) relations between classes.

Example: Ontological Model of Money Launder Domain



Ontological Modeling with XLink

Concept	Example
Classes	propertyAccountRecord, cancelRequestRecord and creditRequestRecord (ref: figs. 5 and 8)
Instances	An instance of a propertyAccountRecord is one bearing "741032-1071" as the policyAccountIdentifier (ref: fig 2, a.1).
Relations: hasPart, sameAs	The three properties, policyAccountIdentifier, cancelRequestIdentifier, and creditRequestIdentifier are equivalent (sameAs) (ref. figs. 8 and 9).
Properties	policyAccountIdentifier, policyStartDate, policyHolder, policyInsurer are properties of a policyAccountRecord (ref: figs. 5, 8 and 9).
Values	"USD" and "5000.00" are the values of payeeCurrency and premiumAmount respectively for a particular instance of a creditRequestRecord (ref: fig. 2, b.2).
Rules	The three properties, policyAccountIdentifier, cancelRequestIdentifier, and creditRequestIdentifier are equivalent (sameAs) if they evaluate to the same value, for example, "741032-1071".

XLink-XPointer-based Concept Modeling

Attribute	Value	Description
xlink:type	extended	Parent element, which defines a complex link in which multiple links can be combined based on other attributes.
	resource	Child element of extended-Type element, which provides a local resource to participate in the link.
	locator	Child element of extended-Type element, which specifies the location of a remote resource participating in the link.
	arc	Child element of extended-Type element, which define traversal rules between the link's participating resources.
xlink:label	<string>	Traversal attribute of extended-, resource-Type elements, which provides a reference (of itself) to arc-Type in creating a traversal arc.
xlink:from, xlink:to	<string>	Traversal attributes of arc-Type element, which define the source and target resources of the arc link.

XLink-XPointer-based Concept Modeling

Attribute	Value	Description
xlink:href	URL	The linked URL
xlink:role	<string>	Semantic attribute of extended-, resource-Type elements, which indicates a property of the resource in a computer readable-form.
xlink:arcrole	<string>	Semantic attribute of arc-Type element, which coincides with the [RDF] notion of a property, where "the role can be interpreted as stating that "starting-resource HAS arc-role ending-resource."
#xpointer		Creates XPointer fragment links with syntax: #xpointer(id(""))

Example (part 1): Policy Account Schema with XLink-XPointer Syntax

```
1    ...
2    <xs:element name="policyAccountRecord" minOccurs="0"
      maxOccurs="unbounded" dfdl:lengthKind="implicit"
      xlink:label="PolicyAccountRecord" xlink:href="http://tmp1.linuz.org/policyAccountSchema.dfdl">
3    ...
4    <xs:element name="policyAccountIdentifier"
      type="xs:string" dfdl:lengthKind="explicit" dfdl:length="20" xlink:label="
      PolicyAccountIdentifier" xlink:type="resource" xlink:href="http://linuz1/
      policyAccountSchema.dfdl#xpointer(///policyAccountIdentifier[@xs:string=
      value])"/>
5    ...
```

Example (part 2): Linkbase loads on extraction

```
1 <linkbase xmlns:xlink="http://www.w3.org/1999/xlink"
  xlink:linkbase="http://www.w3.org/1999/xlink/
  properties/linkbase">
2   <link xlink:type="extended" xlink:title="
  moneyLaunderLinkbase">
3
4     <!-- Linkbase loads on extraction request. -->
5     <basesloaded>
6       <startsrc xlink:label="filter_spec" xlink:href="/
  local/filter_spec.xml#params" />
7       <linkbase xlink:label="linkbase" xlink:href="/local/
  linkbase.xml" />
8       <load xlink:from="filter_spec" xlink:to="linkbase"
  actuate="onRequest" />
9     </basesloaded>
10
11     ...
```

Example (part 3): Linkbase arcs

```
1      ...
2
3      <!-- Arcs between linkbase and DFDL-data stores. -->
4      <invokeStoreArc xlink:type="arc" xlink:arcrole="
5          linkbase" xlink:from="linkbase" xlink:from="
6          PolicyAccountRecord"/>
7      <invokeStoreArc xlink:type="arc" xlink:arcrole="
8          linkbase" xlink:from="linkbase" xlink:from="
          CancelRequestRecord"/>
9      <invokeStoreArc xlink:type="arc" xlink:arcrole="
10         linkbase" xlink:from="linkbase" xlink:from="
11         RefundRequestRecord"/>
12     ...
```

Example (part 4): Linkbase locators

```
1      ...
2
3      <!-- Locator elements. -->
4      <loc xlink:type="locator" xlink:label="
        PolicyAccountRecord" xlink:href="http://tmp1.
        linuz.org/policyAccountSchema.dfdl"/>
5      <loc xlink:type="locator" xlink:label="
        PolicyAccountIdentifier" xlink:href="http://tmp1
        .linuz.orglinuz1/policyAccountSchema.dfdl#
        xpointer(/////policyAccountIdentifier[@xs:string=
        value])"/>
6      <loc xlink:type="locator" xlink:label="
        PolicyStartDate" xlink:href="http://tmp1.linuz.
        org/policyAccountSchema.dfdl#xpointer(/////
        policyStartDate[@xs:date=value])"/>
7      ...
```

Example (part 5): Linkbase relationship between insurance records

```
1      ...
2
3      <!-- Relationship between policy account, cancel
4           request and refund request identifiers. -->
5      <invokeldArc xlink:type="arc" xlink:arcrole="
6           owl:sameAs" xlink:from="PolicyAccountIdentifier"
7           xlink:to="CancelRequestIdentifier"/>
8      <invokeldArc xlink:type="arc" xlink:arcrole="
           owl:sameAs" xlink:from="PolicyAccountIdentifier"
           xlink:to="RefundRequestIdentifier"/>
9      <invokeldArc xlink:type="arc" xlink:arcrole="
           owl:sameAs" xlink:from="CancelRequestIdentifier"
           xlink:to="PolicyAccountIdentifier"/>
10     ...
```

Example (part 6): Linkbase relationship between parts

```
1      ...
2      <!-- Relationship between policy account record
3           and its parts. -->
4      <invokeParArc xlink:type="arc" xlink:arcrole="
5           owl:hasPart" xlink:from="PolicyAccountRecord"
6           xlink:to= "PolicyAccountIdentifier" />
7      <invokeParArc xlink:type="arc" xlink:arcrole="
8           owl:hasPart" xlink:from="PolicyAccountRecord"
9           xlink:to= "PolicyStartDate" />
10     <invokeParArc xlink:type="arc" xlink:arcrole="
11          owl:hasPart" xlink:from="PolicyAccountRecord"
12          xlink:to= "PolicyHolder" />
13     <invokeParArc xlink:type="arc" xlink:arcrole="
14          owl:hasPart" xlink:from="PolicyAccountRecord"
15          xlink:to= "PolicyInsurer" />
16     ...
```

- For syntactic analysis, construct a DFDL grammar as a set of recursive definitions. This effort entails creating a context free grammar (CFG) production rule for each code fragment in the DFDL schema language.
- For semantic analysis:
 - Devise the desired higher-order abstract syntax (HOAS) for DFDL. This also entails creating a HOAS code fragment for each CFG production rule.
 - Devise CRS transformation rules to address a multitude of issues such as DFDL scoping semantics. DFDL addresses scoping semantics through transformation rules by exploiting lambda lifting and dropping

What is CRSX?

- Rewriting is theory of stepwise or discrete transformations of objects
- Combinatory Reduction System (CRS) extends the format of first-order rewriting (TRS) with an operator for abstraction over variables
- A legal TRS cannot express and manipulate terms with bound variables,
 - e.g. untyped λ -calculus with a beta reduction,
$$(\lambda x.M)N \rightarrow M[x:=N]$$
 - e.g. a map function that applies a function to all elements of a list,
$$\text{map}(f,\text{nil}) \rightarrow \text{nil}$$
$$\text{map}(f, \text{cons}(h,t)) \rightarrow \text{cons}(f(h),\text{map}(f,t))$$
- Kristoffer Rose contributes CRSX, which implements Klop's CRS with extensions to support the writing of compilers

Example 1 —CRSX Rewriting Rule System

—Consider example of the rewriting rule system,

+ [3,1]

PlusS: + [S[#1],#2] \rightarrow S [+ [#1,#2]];

—Consider the meaning of the previous syntax [6]:

—Consider the form: **name[options]:pattern** \rightarrow **contraction**, where name should be a constructor and the pattern and contraction should be terms

—Capitalized words (e.g. Cons and Nil) , numbers, and most other symbols are **constructors**, which take an optional ordered or positional **parameter** list in immediately following []s

—Each parameter is itself a **term**, and called a **subterm**

—Names containing # are special pattern variables, or **meta-variables**, that are used to match arbitrary subterms at the indicated position

Example 2 —CRSX Rewriting Rule System

—Consider example of the rewriting rule system,

Let[E1, x.E2]

Let[Copy[#1]]: Let[#1, x.#2[x]] \rightarrow #2[#1]

—Consider the meaning of the previous syntax [6]:

—Uncapitalized words (e.g. x and foo) denote **variables**

—Variables are useful for defining binding constructs. It allows **explicit scoping** consider,

e.g. $E ::= \text{let } x := E1 \text{ in } E2$

—Let[#1, x.#2[x]], where #1 is a **meta-variable** pattern and x.#2[x] is a **meta-application** pattern; the constructor must be Let; there are exactly two subterms; the second subterm under the binder can be anything wherein the variable matched by x may occur, and is matched by #₂ where we keep track of all the actual **occurrences** of the bound variable

Example (part 1): Syntax to CFG Production Rule for DFDL

```
1 <xs:annotation>
2   <dfdl:defineEscapeScheme name= "myEscapeScheme">
3     ...
4   </dfdl:escapeScheme escapeCharacter= "/" />
5     ...
6 <dfdl:element representation="text" escapeSchemeRef="
7   myEscapeScheme"
```

Figure: Consider a small fragment of the DFDL Language that allows us to construct a data description of an *escape character*, which is used to invoke an alternative interpretation on subsequent characters in a character sequence.

Modeling of Logical Datatypes & Constraints

Logical Datatypes & Constraints	Context Free Grammar (CFG) example	Higher-Order Abstract Syntax (HOAS) example
Structures (xs:complexType)	<pre>2 <decl_xs> ::= "<" XS_COMPONENT <stmt>* ">/" XS_COMPONENT ">" 11 XS_COMPONENT ::= "xs:complexType"</pre>	XsComponent[ComplexType]
Atomic data values (xs:simpleType)	<pre>11 XS_COMPONENT ::= "xs:simpleType"</pre>	XsComponent[SimpleType]
Ordering (xs:sequence or xs:choice)	<pre>11 XS_COMPONENT ::= "xs:sequence" "xs:choice"</pre>	XsComponent[Sequence], XsComponent[Choice]
Occurrences (xs:minOccurs or xs:maxOccurs)	<pre>5 <stmt> ::= XS_ATTRIBUTE "=" <xs_attribute_value> 16 XS_ATTRIBUTE ::= "xs:minOccurs" "xs:maxOccurs" 8 <xs_attribute_value> ::= <xs_number></pre>	ComponentAttribute[MinOccurs], ComponentAttribute[MaxOccurs]

Example (part 2): Syntax to CFG Production Rule for DFDL

```
1 2 <dcl_xs> ::= "<" XS_ADMIN <stmt>* ">" <dcl_xs>* "</"  
   XS_ADMIN ">" | "<" XS_COMPONENT <stmt>* ">" <dcl_xs>  
   * "</" XS_COMPONENT ">"  
2  
3 3 <dcl_dfdl> ::= "<" DFDL_ADMIN <stmt>* ">" <dcl_dfdl> "  
   </" DFDL_ADMIN ">" | "<" DFDL_COMPONENT <stmt>* ">"  
   <dcl_dfdl> "</" DFDL_COMPONENT ">" | ...  
4  
5 5 <stmt>      ::= "escapeSchemeRef" = <NCName_value> |  
   "name" = <QName> | ...  
6  
7 15 DFDL_ADMIN ::= "dfdl:defineEscapeScheme" | ...
```

Figure: The following production rules specify how these data descriptions are formed:

Example (part 3): CFG Production Rule to HOAS for DFDL

```
1 TERM ::=(  
2   Let[ VALUE, TYPE, x::VALUE . TERM ];  
3   Lam[ VALUE, TYPE, x::VALUE . TERM ];  
4   Context[ ];  
5   Element[KIND, $List[ATTRIBUTE], $List[DFDL_PROPERTY],  
6           $List[XLP_ATTRIBUTE], TERM];  
7   Pair[ TERM, TERM ];  
8   Nil;  
9   T-Attribute  
10  T-BuildSchema  
11  T-BuildElement  
12  XML-Visit[ XLink-XPointer ]  
13 );
```

Figure: Consider our top level terms for the DFDL CRSX system after normalization. Consider our top level term for representing explicit scoping. The terms are written in the form of a higher-order abstract syntax.

Example (part 4): Using rewrite rules to define explicit scoping of attributes

```
1 XsComponent-Attribute [ Copy[#QName] ]
2 :
3 {#Env; #QName: ComponentAttribute[#kind]}
4 XML-Attribute [ #prefix , #QName, #Value , ok.#Continuation [
      ok ] ]
5 ->
6 {#Env}
7 Let [ #Value , a.{#Env} AddXsAssoc[#prefix , #QName, a , ok.#
      Continuation [ok] ]
8 ;
```

Figure: The following rewrite rule match "XML-Attribute" constructor with parameters and appends the entry to the the local environment:

Example (part 5): Using rewrite rules to define nested scopes of elements

```
1  
2 <dfdl:discriminator ...> ... </dfdl:discriminator>
```

Figure: Consider the block scope a DFDL element

Example (part 6): Using rewrite rules to define nested scopes of elements

```
1 DfdIValidation-Start [Copy[#QName]]
2 :
3 {#Env; #QName: DfdIValidation[#kind]}
4 XML-Start[#prefix, #QName, ok prefix.#Continuation[ok,
   prefix]]
5 —>
6 {#Env; "T-BuildElement": Save-BuildElement[#QName]} #
   Continuation[OK, #QName]
7 ;
8 ...
```

Figure: The following rewrite rules "mark" the start-of block in the environment:

Example (part 7): Using rewrite rules to define nested scopes of elements

```
1 DfdIValidation-End
2 :
3 {#Env}
4 XML-End[ #prefix , ok.#Continuation [ok]]
5 —>
6 {#Env}
7 T-BuildSchema [OK, #QName, DfdIValidation[#kind], #content ,
  ok.#[ok]]
```

Figure: The following rewrite rules "mark" the end-of block in the environment:

Example (part 8): Using rewrite rules to fold and unfold the DFDL program

```
1 1 <dcl_schema> ::= "<" XS_SCHEMA <stmt>* ">" <dcl_xs>*  
   "</" XS_SCHEMA">" $  
2  
3 3 <dcl_xs> ::= "<" DFDL_COMPONENT <stmt>* ">" <dcl_xs>  
   "</" DFDL_COMPONENT ">"
```

Figure: Consider the recursive definition of Schema element:







Example (part 9): Using rewrite rules to fold and unfold the DFDL program

```
1 Schema-End
2 :
3 {#Env}
4 XML-End[ #prefix , ok.#Continuation [ok]]
5 →
6 {#Env}
7 T-BuildSchema [OK, #QName, Schema , #content , ok.#[ok]]
8     ...
9 -[Copy[#QName]]
10 :
11 {#Env}
12 T-BuildSchema [OK, #QName, Schema , #content , ok.#[ok]]
13 →
14 Element[#QName, Schema , ${#{#Env}Get , "ATTRIBUTE" , ()} , [
    ${#{#Env}Get , "DFDL_PROPERTY" , ()} , ${#{#Env}Get , "
    CONTENT" , ()} ] ]
15 ;
```

The future work includes:

- Specifying the transformation and evaluation of the DFDL/XLink/HOF specification into parser combinator form
- Investigating the operational semantics of the higher-order function (HOF) and linking abstractions in order to optimize distributed data extraction
- Generating comparative performance metrics

References

-  F. A. T. Force. (2014 (accessed August 19, 2014)) Money laundering.
-  I. A. of Insurance Supervisors, *Examples of Money Laundering and Suspicious Transactions Involving Insurance*. International Association of Insurance Supervisors, 2004. [Online]. Available: <http://books.google.com/books?id=bSvoHAAACAAJ>
-  O. D. WG, S. M. Hanson, and A. W. Powell, "Data format description language (dfdl) v1. 0 specification."
-  S. DeRose, E. Maler, D. Orchard, and N. Walsh, "Xml linking language (xlink) version 1.1, w3c recommendation 06 may 2010," 2010.
-  P. Grosso, E. Maler, J. Marsh, and N. Walsh, "Xpointer framework," *W3c recommendation*, vol. 25, 2003.
-  K. H. Rose, "Crsx-combinatory reduction systems with extensions," in *LIPICs-Leibniz International Proceedings in Informatics*, vol. 10. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2011.

The End