

Data Analytics and Visualization Environment  
for xAPI and the Total Learning Architecture:  
DAVE Learning Analytics Algorithms

Yet Analytics

September 16, 2019

# Introduction

This report introduces the updated definition of learning analytics algorithms in terms of **Operations**, **Primitives** and **Algorithms** and presents an updated definition for each of the previously defined algorithms. The previous definitions will be included for reference. In a more general sense, this report establishes a set of style guidelines for the reporting of algorithms and associated visualization templates.

This document will be updated to include additional Operations, Primitives and Algorithms as they are defined by the Author of this report or members of the Open Source Community. Updates may also address refinement of existing definitions and this document should be understood to be an example of algorithm presentation and not the final state of any defined algorithm.

The structure of this documents is as follows:

1. A formal specification for xAPI written in Z
2. An Introduction to Terminology of Operations, Primitives and Algorithms
3. What is an Operation
4. What is a Primitive
5. What is an Algorithm
6. Foundational Operations
7. Example Primitives
8. An algorithm definition including
  - (a) Init
  - (b) Relevant?
  - (c) Accept?
  - (d) Step
  - (e) Result
9. Previous Algorithm definitions where each consists of
  - (a) an introduction for the algorithm
  - (b) the structure of the ideal input data
  - (c) how to retrieve input data from an LRS
  - (d) the statement parameters which the algorithm will utilize
  - (e) notices regarding data collected during the 2018 pilot test of the TLA
  - (f) a summary of the algorithm

- (g) the formal specification of the algorithm
- (h) pseudocode representation of the algorithm
- (i) JSONSchema for the output of the algorithm
- (j) a description of the associated visualization
- (k) a prototype of the visualization
- (l) a collection of suggestions describing how the algorithm could be adapted to improve the quality of the visualization prototype

# 1 xAPI Formal Specification

The current formal specification only defines xAPI statements abstractly within the context of Z. A concrete definition for xAPI statements is outside the scope of this document.

## 1.1 Basic Types

$IFI ::= mbox \mid mbox\_sha1sum \mid openid \mid account$

- Type unique to Agents and Groups, The concrete definition of the listed values is outside the scope of this specification

$OBJECTTYPE ::= Agent \mid Group \mid SubStatement \mid StatementRef \mid Activity$

- A type which can be present in all activities as defined by the xAPI specification

$INTERACTIONTYPE ::= true-false \mid choice \mid fill-in \mid long-fill-in \mid matching \mid performance \mid sequencing \mid likert \mid numeric \mid other$

- A type which represents the possible interactionTypes as defined within the xAPI specification

$INTERACTIONCOMPONENT ::= choices \mid scale \mid source \mid target \mid steps$

- A type which represents the possible interaction components as defined within the xAPI specification
- the concrete definition of the listed values is outside the scope of this specification

$CONTEXTTYPES ::= parent \mid grouping \mid category \mid other$

- A type which represents the possible context types as defined within the xAPI specification

$[STATEMENT]$

- Basic type for an xAPI data point

$[AGENT, GROUP]$

- Basic types for Agents and collections of Agents

## 1.2 Id Schema

$Id$ $id : \mathbb{F}_1 \#1$
---------------------------------

- the schema  $Id$  introduces the component  $id$  which is a non-empty, finite set of 1 value

### 1.3 Schemas for Agents, Groups and Actors

<i>Agent</i>	
<i>agent</i> : <i>AGENT</i>	
<i>objectType</i> : <i>OBJECTTYPE</i>	
<i>name</i> : $\mathbb{F}_1 \#1$	
<i>ifi</i> : <i>IFI</i>	
<i>objectType</i> = <i>Agent</i>	
<i>agent</i> = $\{ifi\} \cup \mathbb{P}\{name, objectType\}$	

- The schema *Agent* introduces the component *agent* which is a set consisting of an *ifi* and optionally an *objectType* and/or *name*

<i>Member</i>	
<i>Agent</i>	
<i>member</i> : $\mathbb{F}_1$	
<i>member</i> = $\{a : AGENT \mid \forall a_n : a_i..a_j \bullet i \leq n \leq j \bullet a = agent\}$	

- The schema *Member* introduces the component *member* which is a set of objects *a*, where for every *a* within  $a_0..a_n$ , *a* is an *agent*

<i>Group</i>	
<i>Member</i>	
<i>group</i> : <i>GROUP</i>	
<i>objectType</i> : <i>OBJECTTYPE</i>	
<i>ifi</i> : <i>IFI</i>	
<i>name</i> : $\mathbb{F}_1 \#1$	
<i>objectType</i> = <i>Group</i>	
<i>group</i> = $\{objectType, name, member\} \vee \{objectType, member\} \vee \{objectType, ifi\} \cup \mathbb{P}\{name, member\}$	

- The schema *Group* introduces the component *group* which is of type *GROUP* and is a set of either *objectType* and *member* with optionally *name* or *objectType* and *ifi* with optionally *name* and/or *member*

<i>Actor</i>	
<i>Agent</i>	
<i>Group</i>	
<i>actor</i> : <i>AGENT</i> $\vee$ <i>GROUP</i>	
<i>actor</i> = <i>agent</i> $\vee$ <i>group</i>	

- The schema *Actor* introduces the component *actor* which is either an *agent* or *group*

## 1.4 Verb Schema

<i>Verb</i>	_____
<i>Id</i>	
$display, verb : \mathbb{F}_1$	
$verb = \{id, display\} \vee \{id\}$	

- The schema *Verb* introduces the component *verb* which is a set that consists of either *id* and the non-empty, finite set *display* or just *id*

## 1.5 Object Schema

<i>Extensions</i>	_____
$extensions, extensionVal : \mathbb{F}_1$	
$extensionId : \mathbb{F}_1 \# 1$	
$extensions = \{e : (extensionId, extensionVal) \mid \forall e_n : e_i..e_j \bullet i \leq n \leq j \bullet$ $(extensionId_i, extensionVal_i) \vee (extensionId_i, extensionVal_j) \wedge$ $(extensionId_j, extensionVal_i) \vee (extensionId_j, extensionVal_j) \wedge$ $extensionId_i \neq extensionId_j\}$	

- The schema *Extensions* introduces the component *extensions* which is a non-empty, finite set that consists of ordered pairs of *extensionId* and *extensionVal*. Different *extensionIds* can have the same *extensionVal* but there can not be two identical *extensionId* values
- *extensionId* is a non-empty, finite set with one value
- *extensionVal* is a non-empty, finite set

<i>InteractionActivity</i>	_____
$interactionType : INTERACTIONTYPE$	
$correctResponsePattern : seq_1$	
$interactionComponent : INTERACTIONCOMPONENT$	
$interactionActivity = \{interactionType, correctReponsePattern, interactionComponent\} \vee$ $\{interactionType, correctResponsePattern\}$	

- The schema *InteractionActivity* introduces the component *interactionActivity* which is a set of either *interactionType* and *correctResponsePattern* or *interactionType* and *correctResponsePattern* and *interactionComponent*

<i>Definition</i>
<i>InteractionActivity</i>
<i>Extensions</i>
<i>definition, name, description</i> : $\mathbb{F}_1$
<i>type, moreInfo</i> : $\mathbb{F}_1 \#1$
<i>definition</i> = $\mathbb{P}_1\{name, description, type, moreInfo, extensions, interactionActivity\}$

- The schema *Definition* introduces the component *definition* which is the non-empty, finite power set of *name*, *description*, *type*, *moreInfo* and *extensions*

<i>Object</i>
<i>Id</i>
<i>Definition</i>
<i>Agent</i>
<i>Group</i>
<i>Statement</i>
<i>objectTypeA, objectTypeS, objectTypeSub, objectType</i> : <i>OBJECTTYPE</i>
<i>substatement</i> : <i>STATEMENT</i>
<i>object</i> : $\mathbb{F}_1$
<i>substatement</i> = <i>statement</i>
<i>objectTypeA</i> = <i>Activity</i>
<i>objectTypeS</i> = <i>StatementRef</i>
<i>objectTypeSub</i> = <i>SubStatement</i>
<i>objectType</i> = <i>objectTypeA</i> $\vee$ <i>objectTypeS</i>
<i>object</i> = $\{id\} \vee \{id, objectType\} \vee \{id, objectTypeA, definition\}$ $\vee \{id, definition\} \vee \{agent\} \vee \{group\} \vee \{objectTypeSub, substatement\}$ $\vee \{id, objectTypeA\}$

- The schema *Object* introduces the component *object* which is a non-empty, finite set of either *id*, *id* and *objectType*, *id* and *objectTypeA*, *id* and *objectTypeA* and *definition*, *agent*, *group*, or *substatement*
- The schema *Statement* and the corresponding component *statement* will be defined later on in this specification

## 1.6 Result Schema

<i>Score</i>
<i>score</i> : $\mathbb{F}_1$ <i>scaled, min, max, raw</i> : $\mathbb{Z}$
<i>scaled</i> = $\{n : \mathbb{Z} \mid -1.0 \leq n \leq 1.0\}$ <i>min</i> = $n < \text{max}$ <i>max</i> = $n > \text{min}$ <i>raw</i> = $\{n : \mathbb{Z} \mid \text{min} \leq n \leq \text{max}\}$ <i>score</i> = $\mathbb{P}_1\{\text{scaled}, \text{raw}, \text{min}, \text{max}\}$

- The schema *Score* introduces the component *score* which is the non-empty powerset of *min*, *max*, *raw* and *scaled*

<i>Result</i>
<i>Score</i> <i>Extensions</i> <i>success, completion, response, duration</i> : $\mathbb{F}_1 \# 1$ <i>result</i> : $\mathbb{F}_1$
<i>success</i> = $\{\text{true}\} \vee \{\text{false}\}$ <i>completion</i> = $\{\text{true}\} \vee \{\text{false}\}$ <i>result</i> = $\mathbb{P}_1\{\text{score}, \text{success}, \text{completion}, \text{response}, \text{duration}, \text{extensions}\}$

- The schema *Result* introduces the component *result* which is the non-empty power set of *score*, *success*, *completion*, *response*, *duration* and *extensions*

## 1.7 Context Schema

<i>Instructor</i>
<i>Agent</i> <i>Group</i> <i>instructor</i> : $AGENT \vee GROUP$
<i>instructor</i> = $agent \vee group$

- The schema *Instructor* introduces the component *instructor* which can be ether an *agent* or a *group*

<i>Team</i>
<i>Group</i> <i>team</i> : $GROUP$
<i>team</i> = $group$



- The schema *Team* introduces the component *team* which is a *group*

<i>Context</i> <i>Instructor</i> <i>Team</i> <i>Object</i> <i>Extensions</i> <i>registration, revision, platform, language</i> : $\mathbb{F}_1 \#1$ <i>parentT, groupingT, categoryT, otherT</i> : <i>CONTEXTTYPES</i> <i>contextActivities, statement</i> : $\mathbb{F}_1$
<i>statement</i> = <i>object</i> \ ( <i>id, objectType, agent, group, definition</i> ) <i>parentT</i> = <i>parent</i> <i>groupingT</i> = <i>grouping</i> <i>categoryT</i> = <i>category</i> <i>otherT</i> = <i>other</i> <i>contextActivity</i> = { <i>ca</i> : <i>object</i> \ ( <i>agent, group, objectType, objectTypeSub, substatement</i> )} <i>contextActivityParent</i> = ( <i>parentT, contextActivity</i> ) <i>contextActivityCategory</i> = ( <i>categoryT, contextActivity</i> ) <i>contextActivityGrouping</i> = ( <i>groupingT, contextActivity</i> ) <i>contextActivityOther</i> = ( <i>otherT, contextActivity</i> ) <i>contextActivities</i> = $\mathbb{P}_1$ { <i>contextActivityParent, contextActivityCategory,</i> <i>contextActivityGrouping, contextActivityOther</i> } <i>context</i> = $\mathbb{P}_1$ { <i>registration, instructor, team, contextActivities, revision,</i> <i>platform, language, statement, extensions</i> }

- The schema *Context* introduces the component *context* which is the non-empty powerset of *registration, instructor, team, contextActivities, revision, platform, language, statement* and *extensions*
- The notation *object* \ *agent* represents the component *object* except for its subcomponent *agent*

## 1.8 Timestamp and Stored Schema

<i>Timestamp</i> <i>timestamp</i> : $\mathbb{F}_1 \#1$
<i>Stored</i> <i>stored</i> : $\mathbb{F}_1 \#1$

- The schema *Timestamp* and *stored* introduce the components *timestamp* and *stored* respectively. Each are non-empty, finite sets containing one value

## 1.9 Attachements Schema

<i>Attachments</i>
$display, description, attachment, attachments : \mathbb{F}_1$ $usageType, sha2, fileUrl, contextType : \mathbb{F}_1 \#1$ $length : \mathbb{N}$
$attachment = \{usageType, display, contentType, length, sha2\} \cup \mathbb{P}\{description, fileUrl\}$ $attachments = \{a : attachment\}$

- The schema *Attachments* introduces the component *attachments* which is a non-empty, finite set of the component *attachment*
- The component *attachment* is a non-empty, finite set of the components *usageType*, *display*, *contentType*, *length*, *sha2* with optionally *description* and/or *fileUrl*

## 1.10 Statement and Statements Schema

<i>Statement</i>
$Id$ $Actor$ $Verb$ $Object$ $Result$ $Context$ $Timestamp$ $Stored$ $Attachments$ $statement : STATEMENT$
$statement = \{actor, verb, object, stored\} \cup$ $\mathbb{P}\{id, result, context, timestamp, attachments\}$

- The schema *Statement* introduces the component *statement* which consists of the components *actor*, *verb*, *object* and *stored* and the optional components *id*, *result*, *context*, *timestamp*, and/or *attachments*
- The schema *Statement* allows for subcomponent of *statement* to be referenced via the . (selection) operator

<i>Statements</i>
$Statement$ $IsoToUnix$ $statements : \mathbb{F}_1$
$statements = \{s : statement \mid \forall s_n : s_i..s_j \bullet i \leq n \leq j$ $\bullet convert(s_i.timestamp) \leq convert(s_j.timestamp)\}$

- The schema *Statements* introduces the component *statements* which is a non-empty, finite set of the component *statement* which are in chronological order.

## 2 Operations, Primitives and Algorithms

The following sections introduce, define and explain Operations, Primitives and Algorithms generally using the Terminology presented below. Operations are the building blocks of Primitives whereas Primitives are the building blocks of Algorithms. The definitions which follow are flexible enough to support implementation across programming languages but have been inspired by the core concepts found within Lisp. The focus of these sections is to define the properties of and interactions between Operations, Primitives and Algorithms in a general way which doesn't place unnecessary bounds on their range of possible functionality with respect to processing xAPI data.

### 2.1 Terminology

In the subsections and sections which follow, (s) indicates one or more

#### 2.1.1 Scalar

Singular value  $x$  of a fundamental JSON type as described by [JSON Schema](#)

#### 2.1.2 Collection

a n-tuple of items  $x$  such that

$$X = \langle x_i..x_n..x_j \rangle$$

where

$$i \leq n \leq j \Rightarrow i \prec n \prec j \iff i \neq n \neq j$$

#### 2.1.3 Key

A lookup  $k$  for a  $v$  within a  $kv$  where  $k = x \vee X$

#### 2.1.4 Value

a piece of data  $v$  where  $v = x \vee X$

#### 2.1.5 Key Value pair

Association between a  $k$  and  $v$  where

$$k \mapsto v$$

such that

$$kv = k \mapsto v$$

and a collection of Key Value pair(s) is defined as

$$KV = \langle kv_i..kv_n..kv_j \rangle$$

such that

$$k_n \mapsto v_n$$

and all  $k$  within  $KV$  are unique

$$i_k \neq n_k \neq j_k$$

but the same is not true for all  $v$  within  $KV$

$$i_v = n_v \vee i_v \neq n_v \quad i_v = j_v \vee i_v \neq j_v \quad j_v = n_v \vee j_v \neq n_v$$

### 2.1.6 Statement

Immutable collection of Key Value Pair(s) conforming to the [xAPI Specification](#) as described in the previous section

### 2.1.7 Algorithm State

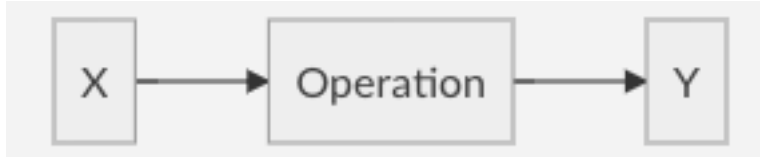
Mutable collection of Key Value Pair(s)

### 2.1.8 Option

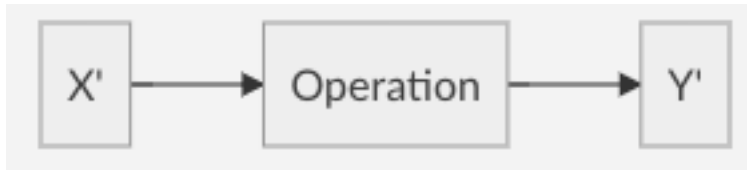
Collection of Key Value Pair(s) which alter the result of an Algorithm

### 3 Operation

Given an input X, an Operation produces output Y



If X changes to X' then the Operation results in Y' instead of Y



#### 3.1 Domain

Any of the following

- Key(s)
- Value(s)
- Key Value pair(s)
- Statement(s)
- Algorithm State

#### 3.2 Range

Any of the following dependent upon the Domain and Functionality of the Operation

- Key(s)
- Value(s)
- Key Value pair(s)
- Statement(s)
- Algorithm State

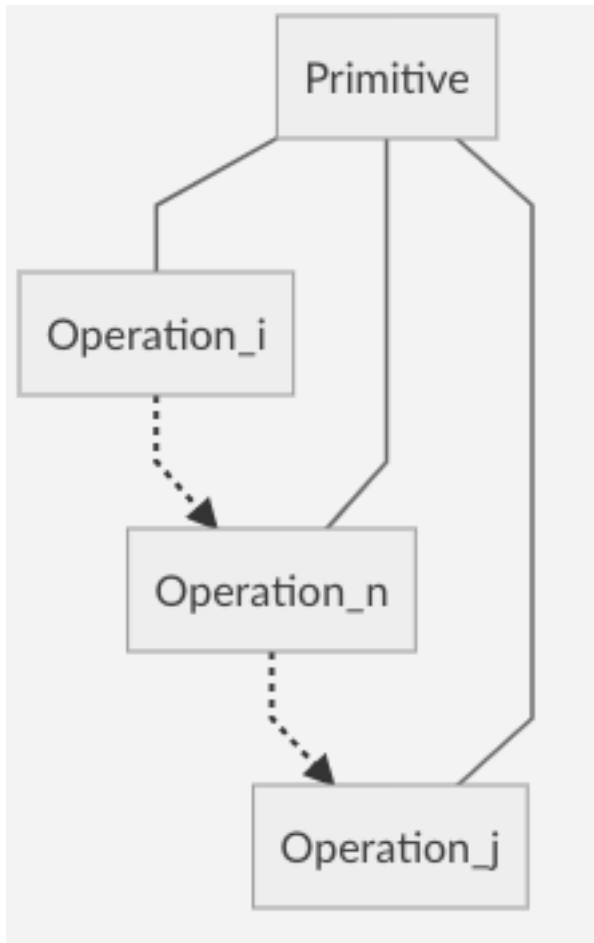
### 3.3 Formal Definition

A relationship between input and output data which will result in the same  $Y$  given the same  $X$

$$\begin{aligned} & Operation(X) = Y \wedge Operation(X') = Y' \\ & \Rightarrow \\ & Y = Y' \iff X = X' \wedge Y \neq Y' \iff X \neq X' \end{aligned}$$

## 4 Primitive

A collection of Operations where the output of an Operation is passed as the argument to the next Operation



Primitives break the processing of xAPI data down into discrete units that can be composed to create new analytical functions. Primitives allow users to address the methodology of answering research questions as a sequence of generic algorithmic steps which establish the necessary data transformations, aggregations and calculations required to reach the solution in an implementation agnostic way.



## 4.1 Domain

Any of the following

- Key(s)
- Value(s)
- Key Value pair(s)
- Statement(s)
- Algorithm State

## 4.2 Range

Any of the following dependent upon the Domain and Functionality of the Primitive

- Key(s)
- Value(s)
- Key Value pair(s)
- Statement(s)
- Algorithm State

## 4.3 Formal Definition

A collection of Operation(s)  $O_n$  labeled  $p$  and defined as

$$p = \langle O_i..O_n..O_j \rangle$$

where

$$i \leq n \leq j \Rightarrow i \prec n \prec j \iff i \neq n \neq j$$

such that the output  $Z$  is defined as the sequential composition of operation(s)  $O_n$  given arg(s)  $Args$  provided to  $p$

$$Z = p(Args) = O_j(O_n(O_i(Args)))$$

## 5 Algorithm

Given a collection of statement(s)  $S_{<a..b..c>}$  and potentially option(s)  $opt$  and potentially an existing Algorithm State  $state$  an Algorithm  $A$  executes as follows

1. call *init*
2. for each  $stmt \in S_{<a..b..c>}$ 
  - (a) *relevant?*
  - (b) *accept?*
  - (c) *step*
3. return *result*

with each process within  $A$  is enumerated as

```
(init [state] body)
- init state

(relevant? [state statement] body)
- is the statement valid for use in algorithm?

(accept? [state statement] body)
- can the algorithm consider the current statement?

(step [state statement] body)
- processing per statement
- can result in a modified state

(result [state] body)
- return without option(s) provided
- possibly sets default option(s)

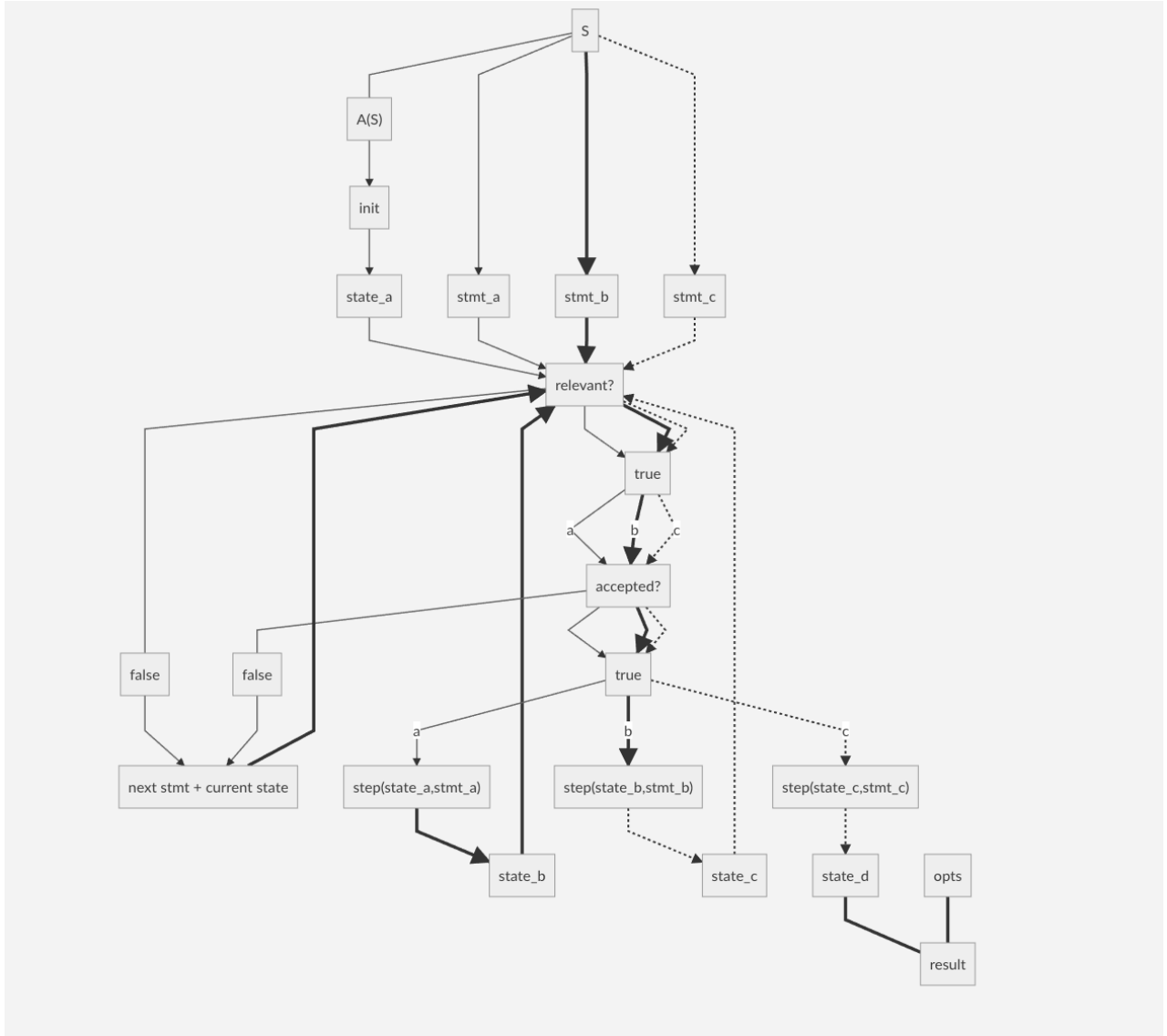
(result [state opts] body)
- return with consideration to option(s)
```

where

- *body* is a collection of Primitive(s)  $P$  which establishes the processing of inputs  $\rightarrow$  outputs
- *state* is a mutable collection of key value pair(s)  $KV$  and synonymous with Algorithm State
- *statement* is a single statement  $s$  within the collection of statements  $S$  passed as input data to the Algorithm  $A$

- *opts* are additional arguments passed to the algorithm *A* which impact the return value of the algorithm

Such that the execution of *A* can be described visually but not exhaustively as



## 5.1 Domain

Any of the following

- Statement(s)
- Algorithm State
- Option(s)

## 5.2 Range

- Algorithm State

## 5.3 Initialization

First process to run within an Algorithm which returns the starting Algorithm State  $state_0$

$$init() = init(state) \vee init() \neq init(state)$$

where  $state_0$  does not need to be related to its arguments

$$init() \rightarrow state_0$$

but  $state_0$  can be derived from some other  $state$  passed as an argument to  $init$

$$init(state) \rightarrow state'_0$$

such that

$$state_0 \neq state'_0$$

but this functionality is dependent upon the *body* of an Algorithms' *init*

### 5.3.1 Domain

- Algorithm State

### 5.3.2 Range

- Algorithm State

## 5.4 Relevant?

First process that each *stmt* passes through such that

$$relevant? \prec accept? \prec step$$

resulting in an indication of whether the *stmt* is valid for use within algorithm *A*

$$relevant?(state, stmt) = true \vee false$$

The criteria which determines validity of *stmt* within *A* is defined by the *body* of *relevant?*

### 5.4.1 Domain

- Statement
- Algorithm State

### 5.4.2 Range

- Scalar

## 5.5 Accept?

Second process that each *stmt* passes through such that

$$relevant? \prec accept? \prec step$$

resulting in an indication of whether the *stmt* can be sent to *step* given the current *state*

$$accept?(state, stmt) = true \vee false$$

The criteria which determines usability of *stmt* given *state* is defined by the *body* of *accept?*

### 5.5.1 Domain

- Statement
- Algorithm State

### 5.5.2 Range

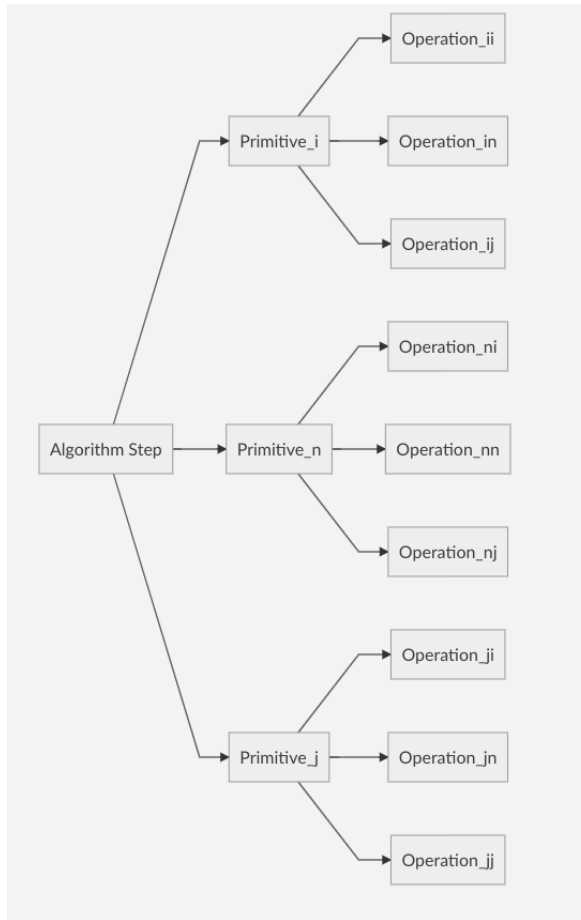
- Scalar

## 5.6 Step

An Algorithm Step consists of a collection of Primitive(s) and therefore collection(s) of Operation(s)



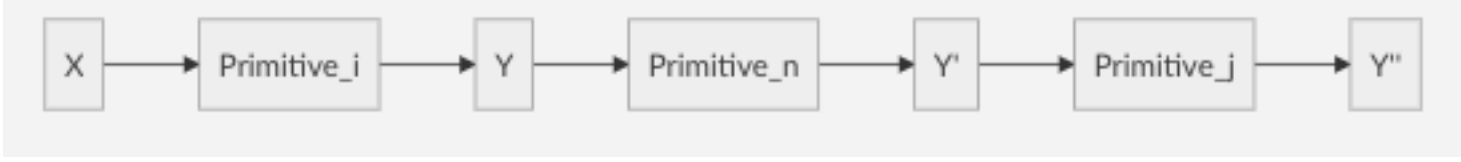
which expands to



$$i \leq n \leq j \Rightarrow i \prec n \prec j$$

$$i_i \leq i_n \leq i_j \leq n_i \leq n_n \leq n_j \leq j_i \leq j_n \leq j_j \Rightarrow i_{<i..n..j>} \prec n_{<i..n..j>} \prec j_{<i..n..j>}$$

where the output of a Primitive is passed as the argument to the next Primitive



The selection and ordering of Operation(s) and Primitive(s) into an Algorithmic Step determines how the Algorithm State changes during iteration through Statement(s) passed as input to the Algorithm.

#### 5.6.1 Domain

- Statement
- Algorithm State

#### 5.6.2 Range

- Algorithm State

#### 5.6.3 Formal Definition

A collection of Primitive(s)

$$P = \langle p_i..p_n..p_j \rangle$$

where

$$i \leq n \leq j \Rightarrow i \prec n \prec j \iff i \neq n \neq j$$

and

$$Z_i = p_i(Args) \Rightarrow O_{ij}(O_{in}(O_{ii}(Args)))$$

where

$$ii \leq in \leq ij \Rightarrow ii \prec in \prec ij \iff ii \neq in \neq ij$$

such that for each  $stmt_b$  within a collection of Statement(s)  $S$  defined as

$$S = \langle stmt_a..stmt_b..stmt_c \rangle$$

where

$$a \leq b \leq c \Rightarrow a \prec b \prec c \iff a \neq b \neq c$$

and

$$a \not\vdash i \wedge b \not\vdash n \wedge c \not\vdash j$$

The output of *step* given a  $stmt_b$  and  $state_b$  is defined as

$$step(state_b, stmt_b) = p_j(p_n(Z_{ib}))$$

where

$$Z_{ib} = p_i(Args) \Rightarrow p_i(state_b, stmt_b) \Rightarrow O_{ij}(O_{in}(O_{ii}(state_b, stmt_b)))$$

and subsequently

$$Z_{nb} = p_n(Z_{ib})$$

which establishes that

$$Z_{jb} = p_j(Z_{nb}) \Rightarrow p_j(p_n(p_i(state_b, stmt_b)))$$

such that for a given  $stmt_b$ ,  $P_{<i..n..j>}$  will always result in a  $Z_{jb}$  but

$$Z_{ib} = state_b \vee state'_{ib} \iff state_b \neq state'_{ib}$$

which means

$$Z_{nb} = p_n(state_b, stmt_b) \vee p_n(state'_{ib}, stmt_b)$$

$$\Rightarrow$$

$$Z_{nb} = state_b \vee state'_{ib} \vee state'_{nb}$$

$$\Rightarrow$$

$$Z_{nb} = Z_{ib} \vee state'_{nb} \iff state_b \neq state'_{ib} \neq state'_{nb}$$

and concludes with

$$Z_{jb} = p_j(state_b, stmt_b) \vee p_j(state'_{ib}, stmt_b) \vee p_j(state'_{nb}, stmt_b)$$

$$\Rightarrow$$

$$Z_{jb} = state_b \vee state'_{ib} \vee state'_{nb} \vee state'_{jb}$$

$$\Rightarrow$$

$$Z_{jb} = Z_{nb} \vee state'_{jb} \iff state_b \neq state'_{ib} \neq state'_{nb} \neq state'_{jb}$$



such that

$$\begin{aligned}
Z_{jb} &\equiv state'_b \\
&\Rightarrow \\
state'_b &= state_b \vee state'_{ib} \vee state'_{nb} \vee state'_{jb} \iff state_b \neq state'_{ib} \neq state'_{nb} \neq state'_{jb}
\end{aligned}$$

the impact being that iteration through all  $stmt \in S < a..b..c >$  results in a return of  $Z_{jc}$  such that

$$Z_{ja} = step(state_a, stmt_a) \Rightarrow state'_a \equiv Z_{ja} = state_a \vee state'_{ia} \vee state'_{na} \vee state'_{ja}$$

and

$$Z_{jb} = step(Z_{ja}, stmt_b) \Rightarrow state'_b \equiv Z_{jb} = Z_{ja} \vee state'_{ib} \vee state'_{nb} \vee state'_{jb}$$

meaning

$$Z_{jc} = step(Z_{jb}, stmt_c) \Rightarrow state'_c \equiv Z_{jc} = Z_{jb} \vee state'_{ic} \vee state'_{nc} \vee state'_{jc}$$

such that each  $stmt \in S < a..b..c >$  may not result in a mutation of  $state$  from  $state \rightarrow state'$

$$\begin{aligned}
state'_c &= Z_{jc} \\
&\Rightarrow \\
state'_c &= state_a \vee state'_{ia} \vee state'_{na} \vee state'_{ja} \vee state'_{ib} \vee state'_{nb} \vee state'_{jb} \vee state'_{ic} \vee state'_{nc} \vee state'_{jc} \\
&\Rightarrow \\
state'_c &= state_a \vee state'_c \neq state_a
\end{aligned}$$

The no-op scenario described above is only a possibility of  $step(state_a, stmt \in S_{<a..b..c>})$  but can be predicted to occur given

- The definition of individual Operations  $O$  which constitute a Primitive  $p$

$$Operation(X) = Y \wedge Operation(X') = Y' \Rightarrow Y = Y' \iff X = X'$$

- The ordering of  $O_{<i..n..j>}$  within  $p$

$$i \prec n \prec j$$

- The Primitive(s)  $p$  chosen for inclusion within  $P_{<i..n..j>}$

$$Z_i = p_i(Args) \Rightarrow O_{ij}(O_{in}(O_{ii}(Args)))$$

$$Z_j = p_j(Args) \Rightarrow O_{jj}(O_{jn}(O_{ji}(Args)))$$

$$\forall Args \exists Z_i = Z_j \iff O_{ij}(O_{in}(O_{ii}(Args))) \equiv O_{jj}(O_{jn}(O_{ji}(Args)))$$

$$<p_i, p_j> \equiv <p_j, p_i> \iff Z_i = Z_j$$

- The ordering of  $p \in P_{i..n..j}$  which implies the ordering of  $O \in p_{<i..n..j>} \in P_{<ii..ij..ni..nj..ji..jj>}$

$$i \prec n \prec j \Rightarrow ii \prec in \prec ij \Rightarrow ii \prec ij \prec ni \prec nj \prec ji \prec jj$$

$$P_{i..n..j} = P_{x..y..z} \Rightarrow <p_i, p_n, p_j> \equiv <p_x, p_y, p_z> \iff p_i \equiv p_x \wedge p_n \equiv p_y \wedge p_j \equiv p_z$$

$$\Rightarrow$$

$$P_{i..n..j} = P_{x..y..z} \iff i \mapsto x \wedge n \mapsto y \wedge j \mapsto z \wedge Z_i = Z_x \wedge Z_n = Z_y \wedge Z_j = Z_z$$

- The Key Value pair(s)  $kv \in stmt \in S_{<a..b..c>}$
- The ordering of Statement(s)  $stmt \in S_{<a..b..c>}$  such that  $a \prec b \prec c$

## 5.7 Return

Last process to run within an Algorithm which returns the Algorithm State *state* without preventing subsequent call of *A*

$$relevant? \prec accept? \prec step \prec result \prec relevant? \iff S \neq \emptyset$$

$$\Rightarrow$$

$$relevant? \prec accept? \prec step \prec result \iff S = \emptyset$$

such that if  $S(t_n) = \emptyset$  and at some future point  $j$  within the timeline  $i..n..j$  this is no longer true  $S(t_j) \neq \emptyset$  then

$$A(state_{n-1}, S(t_{n-1})) = state_n = A(init(), S(t_{n-i})) \iff A(state_n, S(t_n)) = state_n$$

such that the statement(s) added to  $S$  between  $t_i$  and  $t_n$  is

$$S(t_{n-i})$$

and the statement(s) added to  $S$  between  $t_n$  and  $t_j$  be

$$S(t_{j-n})$$

such that

$$S(t_{n-i}) \cup S(t_{j-n}) = S(t_{j-i})$$

which means

$$A(\text{init}(), S(t_{j-i})) = \text{state}_j$$

and establishes that  $A$  can pick up from a previous  $\text{state}_n$  without losing track of its own history.

$$A(\text{result}(\text{state}_n), S(t_{j-n})) = A(\text{init}(), S(t_{j-i})) = \text{state}_j$$

$$\Longleftrightarrow$$

$$\text{result}(\text{state}_n) = A(\text{init}(), S(t_{n-i})) = \text{state}_n$$

Which makes  $A$  capable of taking in some  $S_{\langle i..n..j..\infty \rangle}$  as not all  $s \in S_{\langle i..\infty \rangle}$  have to be considered at once. In other words, the input data does not need to persist across the history of  $A$ , only the effect of  $s$  on  $\text{state}$  must be persisted.

Additionally, the effect of  $\text{opts}$  is determined by the *body* within *result* such that

$$A(\text{result}(\text{state}_n), S(t_{j-n}), \text{opts})$$

$$\equiv$$

$$A(\text{init}(), S(t_{j-i}))$$

$$\equiv$$

$$A(\text{init}(), S(t_{j-i}), \text{opts})$$

$$\equiv$$

$$A(\text{result}(\text{state}_n), S(t_{j-n}))$$

Which implies that  $\text{opts}$  may have an effect on  $\text{state}$  but not in a way which prevents backwards compatibility of  $\text{state}$

### 5.7.1 Domain

- Algorithm State
- Option(s)

### 5.7.2 Range

- Algorithm State

## 6 Foundational Operations

### 6.1 Collections

Operations which expect a Collection of Scalar(s)

$$X = \langle x_i..x_n..x_j \rangle$$

#### 6.1.1 Array?

The operation *array?* will return a boolean which indicates if the passed in argument is a Collection

$$\text{array?}(Arg) = \text{true} \vee \text{false}$$

such that if  $X$  is a collection  $\langle x_0, x_1, x_2, x_3, x_4 \rangle$  where

$$x_0 = 0$$

$$x_1 = \text{foo}$$

$$x_2 = \langle \text{baz}, \text{qux} \rangle$$

$$x_3 = \langle \text{abc} \mapsto 123, \text{def} \mapsto 456 \rangle$$

$$x_4 = \langle \langle \text{ghi} \mapsto 789, \text{jkl} \mapsto 101112 \rangle, \langle \text{ghi} \mapsto 131415, \text{jkl} \mapsto 161718 \rangle \rangle$$

then the following return true

$$\text{array?}(X) = \text{true}$$

$$\text{array?}(x_2) = \text{true}$$

$$\text{array?}(x_4) = \text{true}$$

and the following return false

$$\text{array?}(x_0) = \text{false}$$

$$\text{array?}(x_1) = \text{false}$$

$$\text{array?}(x_3) = \text{false}$$

Indicating that *array?* will return true when

- $Arg$  is a Collection or Scalar(s)

but will return false when

- $Arg$  is a non-array Scalar

which emphasises the difference between

- an *object*

$$x_3 = \langle \text{abc} \mapsto 123, \text{def} \mapsto 456 \rangle = KV$$

- a Collection of *object*(s)

$$x_4 = \langle \langle \text{ghi} \mapsto 789, \text{jkl} \mapsto 101112 \rangle, \langle \text{ghi} \mapsto 131415, \text{jkl} \mapsto 161718 \rangle \rangle = \langle KV_0, KV_1 \rangle$$

### 6.1.2 Append

The operation *append* will return a Collection with a Value added at a specified numeric Index.

$$\text{append}(\text{coll}, v, \text{idx}) \rightarrow X'$$

such that if  $X$  is a collection  $\langle x_0, x_1, x_2 \rangle$  where

$$x_0 = 0$$

$$x_1 = \text{foo}$$

$$x_2 = \langle a, b, c \rangle$$

then the  $v$  will be added to  $X$  to create  $X'$

$$\text{append}(X, \text{bar}, 0) = \langle \text{bar}, 0, \text{foo}, \langle a, b, c \rangle \rangle = X'$$

$$\text{append}(X, \text{bar}, 1) = \langle 0, \text{bar}, \text{foo}, \langle a, b, c \rangle \rangle = X'$$

$$\text{append}(X, \text{bar}, 2) = \langle 0, \text{foo}, \text{bar}, \langle a, b, c \rangle \rangle = X'$$

$$\text{append}(X, \text{bar}, 3) = \langle 0, \text{foo}, \langle a, b, c \rangle, \text{bar} \rangle = X'$$

### 6.1.3 Remove

The operation *remove* will return a Collection minus the Value removed from the specified Numeric Index

$$\text{remove}(\text{coll}, \text{idx}) \rightarrow X'$$

such that if  $X$  is a collection  $\langle x_0, x_1, x_2 \rangle$  where

$$x_0 = 0$$

$$x_1 = \text{foo}$$

$$x_2 = \text{baz}$$

then

$$\text{remove}(X, 0) = \langle \text{foo}, \text{baz} \rangle = X'$$

$$\text{remove}(X, 1) = \langle 0, \text{baz} \rangle = X'$$

$$\text{remove}(X, 2) = \langle 0, \text{foo} \rangle = X'$$

and if  $\text{idx}$  does not exist within  $X$ , *remove* will return  $X$  unaltered

$$\text{remove}(X, 3) = \langle 0, \text{foo}, \text{baz} \rangle = X$$

#### 6.1.4 At Index

The operation *atIndex* will return the Value at

- a numeric Index

$$atIndex(X, idx) \rightarrow v$$

- some depth of numeric indexes

$$atIndex(X, < idx_i..idx_n..idx_j >) \rightarrow v$$

such that if  $X$  is a collection  $< x_0, x_1, x_2 >$  where

$$x_0 = 0$$

$$x_1 = foo$$

$$x_2 = < a, b, c >$$

then

$$atIndex(X, 0) = 0$$

$$atIndex(X, 1) = foo$$

$$atIndex(X, 2) = < a, b, c >$$

$$atIndex(X, < 2, 1 >) = b$$

and if *idx* does not exist within  $X$ , *atIndex* will return the representation of nothingness

$$atIndex(X, 3) = nil$$

$$atIndex(X, < 2, 3 >) = nil$$

## 6.2 Key Value Pairs

Operations which expect a Collection of Key Value pair(s)

$$KV = < k_i v_{k_i} .. k_n v_{k_n} .. k_j v_{k_j} >$$

### 6.2.1 Object?

The operation *object?* will return a boolean which indicates if the passed in argument is a mapping of Key(s) to Value(s)

$$object?(Arg) = true \vee false$$

such that if  $KV$  is a collection of Key Value pair(s)  $\langle k_0v_{k_0}, k_1v_{k_1}, k_2v_{k_2}, k_3v_{k_3} \rangle$  where

$$k_0 = abc \wedge v_{k_0} = 123$$

$$\Rightarrow$$

$$k_0v_{k_0} = abc \mapsto 123$$

and

$$k_1 = def \wedge v_{k_1} = xyz \mapsto 456$$

$$\Rightarrow$$

$$k_1v_{k_1} = def \mapsto xyz \mapsto 456$$

and

$$k_2 = ghi \wedge v_{k_2} = \langle 7, 8, 9 \rangle$$

$$\Rightarrow$$

$$k_2v_{k_2} = ghi \mapsto \langle 7, 8, 9 \rangle$$

and

$$k_3 = k_0v_{k_0} \wedge v_{k_3} = v_{k_2}$$

$$\Rightarrow$$

$$k_3v_{k_3} = \langle \langle abc \mapsto 123 \rangle \mapsto \langle 7, 8, 9 \rangle \rangle$$

then the following will return true

$$object?(KV)$$

$$object?(k_0v_{k_0}) \Rightarrow object?(abc \mapsto 123)$$

$$object?(k_1v_{k_1}) \Rightarrow object?(def \mapsto xyz \mapsto 456)$$

$$object?(k_2v_{k_2}) \Rightarrow object?(ghi \mapsto \langle 7, 8, 9 \rangle)$$

$$object?(k_3v_{k_3}) \Rightarrow object?(\langle \langle abc \mapsto 123 \rangle \mapsto \langle 7, 8, 9 \rangle \rangle)$$

$$object?(v_{k_1}) \Rightarrow object?(xyz \mapsto 456)$$

$$object?(k_3) \Rightarrow object?(abc \mapsto 123)$$

and the following return false

$$object?(k_0) \Rightarrow object?(abc)$$

$$object?(k_1) \Rightarrow object?(def)$$

$$object?(v_{k_0}) \Rightarrow object?(123)$$

$$object?(v_{k_2}) \Rightarrow object?(\langle 7, 8, 9 \rangle)$$

### 6.2.2 Associate

The operation *associate* will return a modified KV' with a mapping of

- Key to Value added at a specified Key.

$$associate(KV, k, v) \rightarrow KV'$$

- Key to Value added at a specified nesting of Key(s).

$$associate(KV, < k_i..k_n..k_j >, v) \rightarrow KV'$$

such that if  $KV$  is a collection of Key Value pair(s)  $< k_0v_{k_0}, k_1v_{k_1} >$  where

$$k_0 = abc \wedge v_{k_0} = 123$$

$$\Rightarrow$$

$$k_0v_{k_0} = abc \mapsto 123$$

and

$$k_1 = def \wedge v_{k_1} = xyz \mapsto 456$$

$$\Rightarrow$$

$$k_1v_{k_1} = def \mapsto xyz \mapsto 456$$

such that

$$KV = < abc \mapsto 123, def \mapsto xyz \mapsto 456 >$$

When  $k$  is a single Key

- $k \notin KV$ ,  $k \mapsto v$  is added to  $KV$  to create  $KV'$

$$associate(KV, baz, foo) = KV'$$

$$=$$

$$< abc \mapsto 123, def \mapsto xyz \mapsto 456, baz \mapsto foo >$$

- $k \in KV$ , the previous mapping is overwritten to create  $KV'$

$$associate(KV, abc, 789) = KV'$$

$$=$$

$$< abc \mapsto 789, def \mapsto xyz \mapsto 456 >$$

$$associate(KV, def, 456) = KV'$$

$$=$$

$$< abc \mapsto 123, def \mapsto 456 >$$



When  $k$  is a Collection of Key(s)  $K = \langle k_i, k_j \rangle$

- $K \notin KV$ ,  $k_i \mapsto k_j \mapsto v$  is added to  $KV$  to create  $KV'$

$$associate(KV, \langle baz, bar \rangle, foo) = KV'$$

=

$$\langle abc \mapsto 123, def \mapsto xyz \mapsto 456, baz \mapsto bar \mapsto foo \rangle$$

- $k_i \in KV \wedge k_j \notin v_{k_i} \wedge object?(v_{k_i}) = false$ , the previous mapping at  $k_i$  is overwritten to create  $KV'$

$$associate(KV, \langle abc, cba \rangle, 789) = KV'$$

=

$$\langle abc \mapsto cba \mapsto 789, def \mapsto xyz \mapsto 456 \rangle$$

- $k_i \in KV \wedge k_j \notin v_{k_i} \wedge object?(v_{k_i}) = true$ ,  $k_j \mapsto v$  is added to  $k_i$

$$associate(KV, \langle def, zyx \rangle, fizbuz) = KV'$$

=

$$\langle abc \mapsto 123, def \mapsto \langle xyz \mapsto 456, zyx \mapsto fizbuz \rangle \rangle$$

- $k_i \in KV \wedge k_j \in v_{k_i} \Rightarrow object?(v_{k_i}) = true$ ,  $v_{k_j}$  is replaced with  $v$

$$associate(KV, \langle def, xyz \rangle, 654) = KV'$$

=

$$\langle abc \mapsto 123, def \mapsto xyz \mapsto 654 \rangle$$

### 6.2.3 Dissociate

The operation *dissociate* will remove some  $k \mapsto v$  from  $KV$  given  $k \in KV$

$$dissociate(KV, k) \rightarrow KV'$$

such that if  $KV$  is a collection of Key Value pair(s)  $\langle k_0 v_{k_0}, k_1 v_{k_1} \rangle$  where

$$k_0 = abc \wedge v_{k_0} = 123$$

$\Rightarrow$

$$k_0 v_{k_0} = abc \mapsto 123$$

and

$$k_1 = def \wedge v_{k_1} = xyz \mapsto 456$$

$\Rightarrow$

$$k_1 v_{k_1} = def \mapsto xyz \mapsto 456$$

such that

$$KV = \langle abc \mapsto 123, def \mapsto xyz \mapsto 456 \rangle$$

Which means *disocciate*( $KV, k$ ) results in

- $KV \neq KV' \iff k \in KV$  where  $k \mapsto v_k$  is removed from  $KV$

$$dissociate(KV, abc) = \langle def \mapsto xyz \mapsto 456 \rangle = KV'$$

$$dissociate(KV, def) = \langle abc \mapsto 123 \rangle = KV'$$

- $KV = KV' \iff k \notin KV$  where nothing is removed from  $KV$

$$dissociate(KV, cba) = \langle abc \mapsto 123, def \mapsto xyz \mapsto 456 \rangle = KV' = KV$$

#### 6.2.4 At Key

The operation *atKey* will return the Value  $v$  at some specified

- Top level Key  $k$  within  $KV$

$$atKey(KV, k) \rightarrow v$$

- Nested location  $K = \langle k_i..k_n..k_j \rangle$  within  $KV$

$$atKey(KV, K) \rightarrow v$$

such that if  $KV$  is a collection of Key Value pair(s)  $\langle k_0 v_{k_0}, k_1 v_{k_1} \rangle$  where

$$k_0 = abc \wedge v_{k_0} = 123$$

$$\Rightarrow$$

$$k_0 v_{k_0} = abc \mapsto 123$$

and

$$k_1 = def \wedge v_{k_1} = xyz \mapsto 456$$

$$\Rightarrow$$

$$k_1 v_{k_1} = def \mapsto xyz \mapsto 456$$

such that

$$KV = \langle abc \mapsto 123, def \mapsto xyz \mapsto 456 \rangle$$

When  $k$  is a single Key

- $k \notin KV$ , *atKey* will return the representation of nothingness

$$atKey(KV, cba) = nil$$

- $k \in KV$ , *atKey* will return  $v_k$

$$atKey(KV, k_0) \Rightarrow atKey(KV, abc) = 123$$

$$atKey(KV, k_1) \Rightarrow atKey(KV, def) = xyz \mapsto 456$$

When  $K$  is a Collection of Key(s)

- $K \notin KV$ ,  $atKey$  will return the representation of nothingness

$$atKey(KV, < cba, 321 >) = nil$$

- $k_i \in KV \wedge k_j \notin KV$ ,  $atKey$  will return the representation of nothingness

$$atKey(KV, < def, abc >) = nil$$

- $k_i \in KV \wedge k_j \in KV$ ,  $atKey$  will return the nested value

$$atKey(KV, < def, xyz >) = 456$$

## 7 Example Primitives

COMING SOON

## Updated Algorithm Definitions

The following are examples of the new way in which Algorithms were defined.

## 8 Rate of Completions

COMING SOON

## **9 Timeline Of Learner Success**

COMING SOON

## 10 Which Assessment Questions are the Most Difficult

COMING SOON



## 11 How Often are Recommendations Followed

COMING SOON

## Previous Algorithm Definitions

The following are examples of the previous way in which Algorithms were defined.

## 12 Rate of Completions

As learners engage in activities supported by a learning ecosystem, they will build up a history of learning experiences. When the digital resources of that learning ecosystem adhere to a framework dedicated to supporting and understanding the learner, such as the Total Learning Architecture (TLA), it becomes possible to retell their learning story through data and data visualization. One important aspect of that story is the rate of completion<sup>1</sup> of the various digital resources within the learning ecosystem.

### 12.1 Ideal Statements

In order to accurately portray the rates of completion, there are a few base requirements of the data produced by a Learning Record Provider (LRP). They are as follows:

- statements describing a learner completing an activity should<sup>2</sup> use the verb *http://adlnet.gov/expapi/verbs/completed*
- statements describing a learner completing an activity should report if the learner was successful or not via *\$.result.success*
- statement describing a learner completing a scored activity should report the learners score via *\$.result.score.raw*, *\$.result.score.min* and *\$.result.score.max*
- activities must be uniquely and consistently identified across all statements
- The time at which a learner completed a learning activity must be recorded
  - The timestamp should contain an appropriate level of specificity.
  - ie. Year, Month, Day, Hour, Minute, Second, Timezone
- statements describing a learner completing an activity should report the amount of time taken to complete the activity via *\$.result.duration*

### 12.2 Input Data Retrieval

How to query an LRS via a GET request to the Statements Resource via curl. The following section contains the appropriate parameters with example values as well as the curl command necessary for making the request.<sup>345</sup>

---

<sup>1</sup> Completion can be defined by the presence of the verb completed or by the presence of *\$.result.completion* set equal to true. In this algorithm, completion is defined by the presence of the verb completed regardless of *\$.result.completion*. This decision affects how statements are retrieved and filtered. In the case where completion is defined by *\$.result.completion*, the query to the LRS would not include the verb parameter and there would need to be a filtering process which looks for the presence of *\$.result.completion = true*

<sup>2</sup> See footnote 4

<sup>3</sup> See footnote 1.

<sup>4</sup> See footnote 2.

<sup>5</sup> See footnote 3.

```

Verb = "verb=http://adlnet.gov/expapi/verbs/completed"

Since = "since=2018-07-20T12:08:47Z"

Until = "until=2018-07-21T12:08:47Z"

Base = "https://example.endpoint/statements?"

endpoint = Base + Verb + "&" + Since + "&" + Until

Auth = Hash generated from basic auth

S = curl -X GET -H "Authorization: Auth"
      -H "Content-Type: application/json"
      -H "X-Experience-API-Version: 1.0.3"
      Endpoint

```

### 12.3 Statement Parameters to Utilize

The statement parameter locations here are written in [JSONPath](#). This notation is also compatible with the xAPI Z notation due to the defined hierarchy of components. Within the Z specifications, a variable name will be used instead of the \$

- *\$.timestamp*
- *\$.object.id*

### 12.4 2018 Pilot TLA Statement Problems

The initial pilot test data supports the core requirements of this algorithm but completion statements only reports completion scores via *\$.result.scaled* instead of *\$.result.score.raw*, *\$.result.score.min* and *\$.result.score.max*.<sup>6</sup> Given that the official 2018 pilot test is scheduled to take place on July 27th, 2018, this section may require updates pending future data review.

### 12.5 Summary

1. Query an LRS via a GET request to the statements endpoint using the parameters verb, since and until.

---

<sup>6</sup> The one potential issue with using scaled score is the calculation of scaled is not strictly defined by the xAPI specification but is instead up to the authors of the LRP. This results in the inability to reliably compare scaled scores across LRPs. If *\$.result.score.raw*, *\$.result.score.min* and *\$.result.score.max* are reported for all questions, it becomes possible to reliably compare scores across LRPs by generating a scaled score in a consistent way.

2. group statements by their  $\$.object.id$
3. select time range unit for use within rate calculation. Will default to day.
4. determine the amount of time between the first and last instance of a  $\$.object.id$  (in seconds) and divide it by the time unit. ie if the unit is minute, you would divide by 60.
5. calculate the rate by dividing the count of a group (2) by the number of time units covered by the statements (4) so that the rate is the number of completions per activity per time unit.

## 12.6 Formal Specification

### 12.6.1 Basic Types

$TIMEUNIT ::= \{second\}|\{minute\}|\{hour\}|\{day\}|\{week\}|\{month\}|\{year\}$

### 12.6.2 System State

$RateOfCompletion$
$Statements$
$S_{completions} : \mathbb{F}_1$
$S_{grouped}, S_{timeunit}, S_{processed} : \mathbb{F}$
$S_{completions} = statements$
$S_{grouped} = \{byId : seq_1 statement\}$
$S_{withRate} = \{byGroup : (seq_1 statement, \mathbb{N})\}$
$S_{processed} = \{rate : (id, \mathbb{N}, TIMEUNIT)\}$

- The set  $S_{completions}$  is a non-empty, finite set and is the component *statements* which contains the results of the query to the LRS.
- The sets  $S_{grouped}$ ,  $S_{withRate}$  and  $S_{processed}$  are all finite sets
- the set  $S_{grouped}$  is a finite set of objects *byId* which are non-empty, finite sequences of the component *statement*
- the set  $S_{withRate}$  is a finite set of objects *byGroup* which are ordered pairs of non-empty, finite sequences of the component *statement* and a natural number
- the set  $S_{processed}$  is a finite set of objects *rate* where each contains the component *id*, a natural number and the type  $TIMEUNIT$

### 12.6.3 Initial System State

$InitRateOfCompletion$ $RateOfCompletion$ $T : TIMEUNIT$	
$S_{completions} \neq \emptyset$ $S_{grouped} = \emptyset$ $S_{withRate} = \emptyset$ $S_{processed} = \emptyset$ $T = \{day\}$	

- The set  $S_{completions}$  is a non-empty set which contains the results of the GET request(s) to the LRS
- The sets  $S_{grouped}$  ,  $S_{withRate}$  and  $S_{processed}$  are all initially empty
- the variable T has the type  $TIMEUNIT$  and the value  $\{day\}$

### 12.6.4 Calculate Rate

$IsoToUnix$ $convert : \mathbb{F}_1 \rightarrow \mathbb{N}\#1$ $c? : \mathbb{F}_1$ $c! : \mathbb{N}\#1$	
$c! = convert(c?)$	

- The schema  $IsoToUnix$  introduces the function  $convert$  which takes in a finit set of one thing (a timestamp) and converts it to a single natural number.
- the purpose of this function is to convert an ISO 8601 timestamp to the Unix epoch. The concrete definition of the conversion is outside the scope of this document
  - The Unix epoch is the number of seconds that have elapsed since January 1, 1970 (midnight UTC/GMT), not counting leap seconds.

<i>CalcRateByUnit</i>	
<i>Statement</i>	
<i>IsoToUnix</i>	
<i>CountPerGroup</i>	
$unit? : TIMEUNIT$	
$s?, s! : \mathbb{F}$	
$r : \mathbb{N}$	
$rate : (\mathbb{F}, TIMEUNIT) \rightarrow \mathbb{F}$	
$unit? = \{second\} \Rightarrow 1 \vee \{minute\} \Rightarrow 60 \vee \{hour\} \Rightarrow 3600 \vee$ $\{day\} \Rightarrow 86400 \vee \{week\} \Rightarrow 604800 \vee$ $\{month\} \Rightarrow 2629743 \vee \{year\} \Rightarrow 31556926$ $s? = \{g : seq_1 statement\}$ $s! = rate(s?, unit?)$ $s! = \{s : (g, r) \mid \forall g_n : g_i..g_j \bullet i \leq n \leq j \bullet \exists s_n : (g_n, r_n) \bullet$ $r_n = count(g_n) \div ((convert(last\ g_n.timestamp) - convert(head\ g_n.timestamp)) \div unit?) \}$	

- The schema *CalcRateByUnit* introduces the function *rate* where the input *s?* is a set of objects *g* which are each a non-empty, finite sequence of statements and the input *unit?* represents a unit of time.
- for every  $g_n$  within the range  $g_i..g_j$ , there exists an associated object  $s_n$  which is an ordered pair of  $(g_n, r_n)$  where  $r_n$  is equal to the number of items within  $g_n$  divided by the number of *unit?*s within the time range of  $last\ g_n.timestamp - head\ g_n.timestamp$
- the output of the function *rate* is *s!*, the set of all  $s_n$

### 12.6.5 Processes Results

$\Delta RateOfCompletion$ $GroupByActivityId$ $CalcRateByUnit$ $grouped, processed, withRate : \mathbb{F}$ $r : \mathbb{N}$ $T? : TIMEUNIT$	
$T? = \{day\}$ $grouped = \emptyset$ $grouped' = group(S_{completions})$ $S'_{grouped} = S_{grouped} \cup grouped'$ $withRate \subseteq S'_{grouped}$ $withRate' = rate(withRate, T?)$ $S'_{withRate} = withRate' \cup S_{withRate}$ $processed \subseteq S'_{withRate}$ $processed' = \{p : (id, r, T?) \mid$ $\quad \text{let } \{processed_i..processed_j\} == \{b_i..b_j\} \bullet$ $\quad \forall b_n : b_i..b_j \bullet i \leq n \leq j \bullet \exists p_n : (id_n, r_n, T?) \bullet$ $\quad id_n = (head(first\ b_n)).object.id \wedge$ $\quad r_n = (second\ b_n)\}$ $S'_{processed} = processed' \cup S_{processed}$	

- The schema *AggergateCompletionStatements* outlines how to calculate the rate of completion per \$.object.id per second|minute|hour|day|week|month|year
  1.  $S'_{grouped}$  is the result of grouping the statements within  $S_{completions}$  by their \$.object.id
  2. The groups from (1) are passed to the function *rate* with the variable  $T?$  which controls the unit of time, ie per day vs per week
  3. the result of (2) is then processed to create a triplet of \$.object.id, rate, unit of time for all unique \$.object.id within  $S_{completions}$

### 12.6.6 Return

$\Xi RateOfCompletion$ $AggergateCompletionStatements$ $S_{processed}! : \mathbb{F}$	
$S_{processed}! = S_{processed}$	

- The return value  $S_{processed}!$  is equal to  $S_{processed}$  after the operation described by *AggergateCompletionStatements*



## 12.7 Pseudocode

---

### Algorithm 1: Rate of Completions

---

```

Input:  $S_{completed}$ ,  $timeUnit$ 
Result:  $ratePerObjTu'$ 
 $context = \{\}$ ;
 $ratePerObjTu = []$ ;
while  $S_{completion} \neq \emptyset$  do
    foreach  $s \in S_{completion}$  do
         $id \leftarrow s.object.id$ ;
         $ts \leftarrow convert(s.timestamp)$ ;
        if  $id \notin context$  then
            do
                 $times = [ts]$ ;
                 $context' \leftarrow \{id : times\}$ ;
                 $S'_{completion} \leftarrow S_{completion} \setminus s$ ;
                recur  $context', S'_{completion}$ ;
            else
                do
                     $times' \leftarrow context.id \hat{\ } ts$ ;
                     $context' \leftarrow \{id : times'\}$ ;
                     $S'_{completion} \leftarrow S_{completion} \setminus s$ ;
                    recur  $context', S'_{completion}$ ;
                end
            end
        end
    end
    end
    foreach  $k \in context'$  do
         $allTs \leftarrow context'.k$ ;
         $totalDuration \leftarrow \max(allTs) - \min(allTs)$ ;
         $totalCount \leftarrow \text{count}(allTs)$ ;
         $rate \leftarrow totalCount \div (totalDuration \div timeUnit)$ ;
         $subVec = [k, rate, timeUnit]$ ;
         $ratePerObjTu' \leftarrow ratePerObjTu \hat{\ } subVec$ ;
        recur  $ratePerObjTu'$ ;
    end
return  $ratePerObjTu'$ 

```

---

- Values from Z schemas are used within this pseudocode
- the result of the algorithm is an array of arrays where each subarray contains a *statement.object.id*, the *rate* and the *timeUnit* used to calculate *rate*.

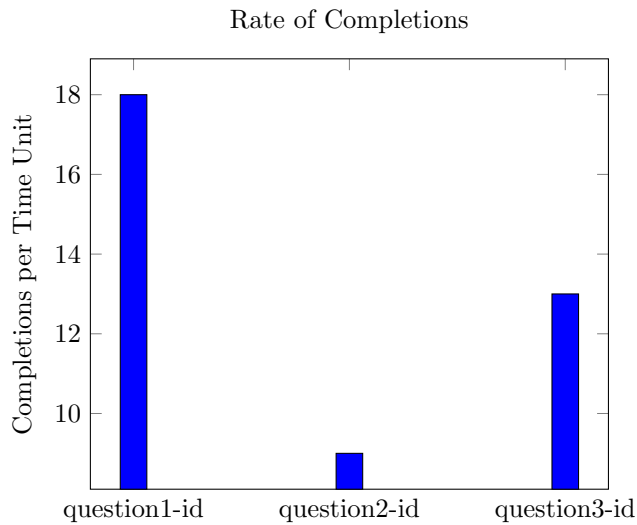
## 12.8 JSON Schema

```
{ "type": "array",  
  "items": { "type": "array",  
    "items": [ { "type": "string" }, { "type": "number" } ],  
  "type": "string" } ] }
```

## 12.9 Visualization Description

The **Rate of Completions** visualization will be a bar chart where the domain consists of *statement.object.id* and the range is a number greater than 0 (the rate of completions for that *statement.object.id*). Every subarray within the array *ratePerObjTu* will be a grouping within the bar chart. The pseudocode specifies an input parameter *timeUnit* which controls the calculation of the rate (range of the visualization). *timeUnit* could be per minute, per day, per week, etc.

## 12.10 Visualization prototype



## 12.11 Prototype Improvement Suggestions

Additional features may be implemented on top of this base specification but they would require adding additional values to each subarray returned by the algorithm. These additional values can be retrieved via (1) performing metadata lookup within or independently of the algorithm (2) by utilizing additional xAPI statement parameters and/or (3) by performing additional computations. The following examples assume the metadata is contained within each statement available to the algorithm.

- use *statement.object.definition.name* instead of *statement.object.id* for x axis label
- populate a tooltip with the people who have completed the activity. This could also include the number of times they have completed it.
- populate a tooltip with the breakdown of which devices or platforms the activity was completed on. This would require the device type or platform to be reported within *statement.context.platform*
- populate a tooltip with the breakdown of percentage successful for all completions of the activity. This would require *statement.result.success*
- populate a tooltip with the breakdown of scores earned (if applicable) for the completions. This would require *statement.result.score.raw*, *statement.result.score.min* and *statement.result.score.max*
- populate a tooltip with the competency associated with the completed activities. The competency should be reported via *statement.context.contextActivities*
- populate a tooltip with the average duration spent to reach completions. This would require *statement.result.duration* to be reported.

## 13 Timeline Of Learner Success

As learners engage in activities supported by a learning ecosystem, they will build up a history of learning experiences. When the digital resources of that learning ecosystem adhere to a framework dedicated to supporting and understanding the learner, such as the Total Learning Architecture (TLA), it becomes possible to retell their learning story through data and data visualization. One important aspect of that story is the learners history of success.

### 13.1 Ideal Statements

In order to accurately portray a learner's timeline of success, there are a few base requirements of the data produced by a Learning Record Provider (LRP). They are as follows:

- the learner must be uniquely and consistently identified across all statements
- learning activities which evaluate a learner's understanding of material must report if the learner was successful or not
  - the grade earned by the learner must be reported
  - the minimum and maximum possible grade must be reported
- The learning activities must be uniquely and consistently identified across all statements
- The time at which a learner completed a learning activity must be recorded
  - The timestamp should contain an appropriate level of specificity.
  - ie. Year, Month, Day, Hour, Minute, Second, Timezone

### 13.2 Input Data Retrieval

How to query an LRS via a GET request to the Statements Resource via curl. The following section contains the appropriate parameters with example values as well as the curl command necessary for making the request.<sup>789</sup>

---

<sup>7</sup>  $S$  is the set of all statements parsed from the statements array within the HTTP response to the Curl request(s). It may be possible that multiple Curl requests are needed to retrieve all query results. If multiple requests are necessary,  $S$  is the result of concatenating the result of each request into a single set

<sup>8</sup> Querying an LRS will not be defined within the following Z specifications but the results of the query will be utilized

<sup>9</sup> If you want to query across the entire history of a LRS, omit Since and Until from the endpoint(s) and remove the associated & symbols.

```

Agent = "agent={\"account\":
    {\"homePage\": \"https://example.homepage\",
      \"name\": 123456}}\"

Since = \"since=2018-07-20T12:08:47Z\"

Until = \"until=2018-07-21T12:08:47Z\"

Base = \"https://example.endpoint/statements?\"

endpoint = Base + Agent + \"&\" + Since + \"&\" + Until

Auth = Hash generated from basic auth

S = curl -X GET -H \"Authorization: Auth\"
-H \"Content-Type: application/json\"
-H \"X-Experience-API-Version: 1.0.3\"
Endpoint

```

### 13.3 Statement Parameters to Utilize

The statement parameter locations here are written in [JSONPath](#). This notation is also compatible with the xAPI Z notation due to the defined hierarchy of components. Within the Z specifications, a variable name will be used instead of the \$

- *\$.timestamp*
- *\$.result.success*
- *\$.result.score.raw*
- *\$.result.score.min*
- *\$.result.score.max*
- *\$.verb.id*

### 13.4 2018 Pilot TLA Statement Problems

The initial pilot test data supports this algorithm. This section may require updates pending future data review following iterations of the TLA testing.

### 13.5 Summary

1. Query an LRS via a [GET](#) request to the statements endpoint using the parameters agent, since and until

2. Filter the results to the set of statements where:
  - $\$.verb.id$  is one of:
    - `http://adlnet.gov/expapi/verbs/passed`
    - `https://w3id.org/xapi/dod-isd/verbs/answered`
    - `http://adlnet.gov/expapi/verbs/completed`
  - $\$.result.success$  is true
3. process the filtered data
  - extract  $\$.timestamp$
  - extract the score values from  $\$.result.score.raw$ ,  $\$.result.score.min$  and  $\$.result.score.max$  and convert them to the scale 0..100
  - create a pair of  $[\$.timestamp, \#]$

## 13.6 Formal Specification

### 13.6.1 Basic Types

$COMPLETION ::=$   
 $\{http : //adlnet.gov/expapi/verbs/passed\} |$   
 $\{https : //w3id.org/xapi/dod - isd/verbs/answered\} |$   
 $\{http : //adlnet.gov/expapi/verbs/completed\}$

$SUCCESS ::= \{true\}$

### 13.6.2 System State

$TimelineLearnerSuccess$	_____
$Statements$	
$S_{all} : \mathbb{F}_1$	
$S_{completion}, S_{success}, S_{processed} : \mathbb{F}$	
$S_{all} = statements$	
$S_{completion} \subseteq S_{all}$	
$S_{success} \subseteq S_{completion}$	
$S_{processed} = \{pair : (statement.timestamp, \mathbb{N})\}$	

- The set  $S_{all}$  is a non-empty, finite set and is the component *statements*
- The sets  $S_{completion}$  and  $S_{success}$  are both finite sets
- the set  $S_{completion}$  is a subset of  $S_{all}$  which may contain every value within  $S_{all}$
- the set  $S_{success}$  is a subset of  $S_{completion}$  which may contain every value within  $S_{completion}$
- the set  $S_{processed}$  is a finite set of pairs where each contains a *statement.timestamp* and a natural number

### 13.6.3 Initial System State

<i>InitTimelineLearnerSuccess</i>	_____
<i>TimelineLearnerSuccess</i>	
$S_{all} \neq \emptyset$	
$S_{completion} = \emptyset$	
$S_{success} = \emptyset$	
$S_{processed} = \emptyset$	

- The set  $S_{all}$  is a non-empty set
- The sets  $S_{completion}$ ,  $S_{success}$  and  $S_{processed}$  are all initially empty

### 13.6.4 Filter for Completion

<i>Completion</i>	_____
<i>Statement</i>	
$completion : STATEMENT \rightarrow \mathbb{F}$	
$s? : STATEMENT$	
$s! : \mathbb{F}$	
$s? = statement$	
$s! = completion(s?)$	
$completion(s?) = \text{if } s?.verb.id : COMPLETION$	
$\quad \text{then } s! = s?$	
$\quad \text{else } s! = \emptyset$	

- The schema *Completion* introduces the function *completion* which takes in the variable  $s?$  and returns the variable  $s!$
- The variable  $s?$  is the component *statement*
- $s!$  is equal to  $s?$  if  $s?.verb.id$  is of the type *COMPLETION* otherwise  $s!$  is an empty set

<i>FilterForCompletion</i>	_____
$\Delta TimelineLearnerSuccess$	
<i>Completion</i>	
$completions : \mathbb{F}$	
$completions \subseteq S_{all}$	
$completions' = \{s : STATEMENT \mid completion(s) \neq \emptyset\}$	
$S'_{completion} = S_{completion} \cup completions'$	

- the set *completions* is a subset of  $S_{all}$  which may contain every value within  $S_{all}$

- The set  $completions'$  is the set of all statements  $s$  where the result of  $completion(s)$  is not an empty set
- the updated set  $S'_{completion}$  is the union of the previous state of set  $S_{completion}$  and the set  $completions'$

### 13.6.5 Filter for Success

$Success$
$Statement$ $success : STATEMENT \rightarrow \mathbb{F}$ $s? : STATEMENT$ $s! : \mathbb{F}$
$s? = statement$ $s! = success(s?)$ $success(s?) = \text{if } s?.result.success : SUCCESS$ $\quad \text{then } s! = s?$ $\quad \text{else } s! = \emptyset$

- the schema  $Success$  introduces the function  $success$  which takes in the variable  $s?$  and returns the variable  $s!$
- the variable  $s?$  is the component  $statement$
- $s!$  is equal to  $s?$  if  $$.result.success$  is of the type  $SUCCESS$  otherwise  $s!$  is an empty set

$FilterForSuccess$
$\Delta TimelineLearnerSuccess$ $Success$ $successes : \mathbb{F}$
$successes \subseteq S_{completion}$ $successes' = \{s : STATEMENT \mid success(s) \neq \emptyset\}$ $S'_{success} = S_{success} \cup successes'$

- the set  $successes$  is a subset of  $S_{completion}$  which may contain every value within  $S_{completion}$
- The set  $successes'$  contains elements  $s$  of type  $STATEMENT$  where  $success(s)$  is not an empty set
- The updated set  $S'_{success}$  is the union of the previous state of  $S_{success}$  and  $successes'$



### 13.6.6 Processes Results

$Scale$ $scaled! : \mathbb{N}$ $raw?, min?, max? : \mathbb{Z}$ $scale : \mathbb{Z} \rightarrow \mathbb{N}$
$scaled! = scale(raw?, min?, max?)$ $scale(raw?, min?, max?) =$ $(raw? * ((0.0 - 100.0) \text{ div } (min? - max?))) +$ $(0.0 - (min? * ((0.0 - 100.0) \text{ div } (min? - max?))))$

- The schema *Scale* introduces the function *scale* which takes 3 arguments, *raw?*, *min?* and *max?*. The function converts *raw?* from the range *min? .. max?* to 0.0..100.0

$ProcessStatements$ $\Delta TimelineLearnerSuccess$ $Scale$ $FilterStatements$ $processed : \mathbb{F}$
$processed \subseteq S_{success}$ $processed' = \{p : (\mathbb{F}_1 \# 1, \mathbb{N}) \mid$ $\quad \text{let } \{processed_i..processed_j\} == \{s_i..s_j\} \bullet$ $\quad i \leq n \leq j \bullet \forall s_n : s_i..s_j \bullet \exists p_n : p_i..p_j \bullet$ $\quad first\ p_n = s_n.timestamp \wedge$ $\quad second\ p_n = scale(s_n.result.score.raw,$ $\quad \quad \quad s_n.result.score.min,$ $\quad \quad \quad s_n.result.score.max)\}$ $S'_{processed} = S_{processed} \cup processed'$

- The operation *ProcessStatements* introduces the variable *processed* which is a subset of  $S_{success}$  which may contain every value within  $S_{success}$
- $S_{success}$  is the result of the operation *FilterStatements*
- The operation defines the variable *processed'* which is a set of objects *p* which are ordered pairs of (1) a finite set containing one value and (2) a single positive number.
- The first component of every object *p*, is the timestamp from the associated *statement* within *processed* ie. *s.timestamp*
- The second component of every object *p* is the result of the function *scale*. The score values contained within the associated *statement* *s* are the arguments passed to *scale*. ie  $scale(s.result.score.raw, s.result.score.min, s.result.score.max)$
- The result of the operation *ProcessStatements* is to updated the set  $S_{processed}$  with the values contained within *processed'*

### 13.6.7 Sequence of Operations

$FilterStatements \hat{=} FilterForCompletion \circ FilterForSuccess$

- The schema  $FilterStatements$  is the sequential composition of operation schemas  $FilterForCompletion$  and  $FilterForSuccess$
- $FilterForCompletion$  happens before  $FilterForSuccess$

$ProcessedStatements \hat{=} FilterStatements \circ ProcessStatements$

- The schema  $ProcessedStatements$  is the sequential composition of operation schemas  $FilterStatements$  and  $ProcessStatements$
- $FilterStatements$  happens before  $ProcessStatements$

### 13.6.8 Return

$Return$
$\exists TimelineLearnerSuccess$
$ProcessedStatements$
$S_{processed}! : \mathbb{F}$
$S_{processed}! = S_{processed}$

- The returned variable  $S_{processed}!$  is equal to the current state of variable  $S_{processed}$  after the operations  $FilterForCompletion$ ,  $FilterForSuccess$  and  $ProcessStatements$

### 13.7 Pseudocode

---

**Algorithm 2:** Timeline of Learner Success

---

```

Input:  $S_{all}$ 
Result:  $coll'$ 
 $coll = []$ ;
while  $S_{all} \neq \emptyset$  do
    foreach  $s \in S_{all}$  do
        if  $s.verb.id = COMPLETION$  then
            do
                 $S'_{completion} \leftarrow s \cup S_{completion}$ ;
                 $S'_{all} \leftarrow S_{all} \setminus s$ ;
                recur  $S'_{completion}, S'_{all}$ ;
            else
                do
                     $S'_{all} \leftarrow S_{all} \setminus s$ ;
                    recur  $S'_{all}$ ;
                end
            end
        end
    end
end
while  $S'_{completion} \neq \emptyset$  do
    foreach  $sc \in S'_{completion}$  do
        if  $sc.result.success = SUCCESS$  then
            do
                 $S'_{success} \leftarrow sc \cup S_{success}$ ;
                 $S'_{completion} \leftarrow S_{completion} \setminus sc$ ;
                recur  $S'_{success}, S'_{completion}$ ;
            else
                do
                     $S'_{completion} \leftarrow S_{completion} \setminus sc$ ;
                    recur  $S'_{completion}$ ;
                end
            end
        end
    end
end
foreach  $ss \in S'_{success}$  do
     $raw? \leftarrow ss.result.score.raw$ ;
     $max? \leftarrow ss.result.score.max$ ;
     $min? \leftarrow ss.result.score.min$ ;
     $scaled \leftarrow scale(raw?, min?, max?)$ ;
     $subVec \leftarrow [ss.timestamp, scaled]$ ;
     $coll' \leftarrow coll \cup subVec$ ;
    recur  $coll'$ 
end
return  $coll'$ 

```

---

- The Z schemas are used within this pseudocode
- The return value `coll` is an array of arrays, each containing a *statement.timestamp* and a scaled score.

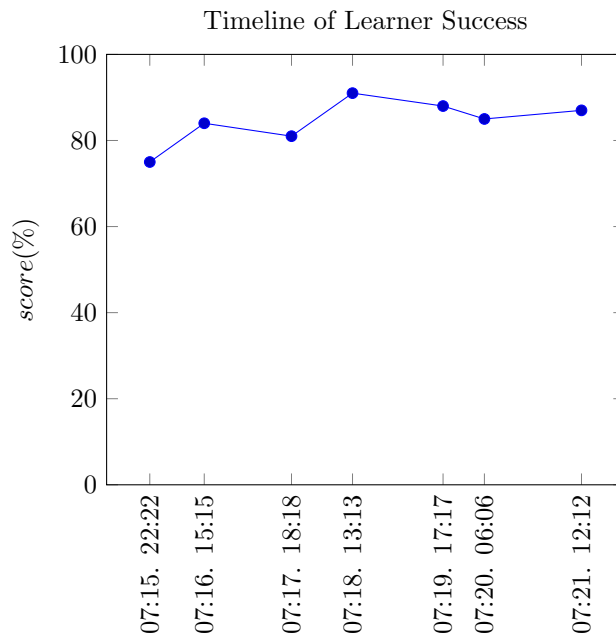
## 13.8 JSON Schema

```
{ "type": "array",
  "items": { "type": "array",
    "items": [ { "type": "string" }, { "type": "number" } ] ] }
```

## 13.9 Visualization Description

The **Timeline of Learner Success** visualization will be a line chart where the domain is time and the range is score on a scale of 0.0 to 100.0. Every subarray will be a point on the chart. The domain of the graph should be in chronological order.

## 13.10 Visualization prototype



## 13.11 Prototype Improvement Suggestions

Additional features may be implemented on top of this base specification but they would require adding additional values to each subarray returned by the

algorithm. These additional values can be retrieved via (1) performing metadata lookup within or independently of the algorithm (2) by utilizing additional xAPI statement parameters and/or (3) by performing additional computations. The following examples assume the metadata is contained within each statement available to the algorithm.

- A tooltip containing the name of an activity when hovering over a specific point on the chart
  - this would require utilizing *\$.object.definition.name*
- A tooltip containing the device on which the activity was experienced
  - this would require utilizing *\$.context.platform*
- A tooltip containing the instructor associated with a particular data point
  - this would require utilizing *\$.context.instructor*

## 14 Which Assessment Questions are the Most Difficult

As learners engage in activities supported by a learning ecosystem, they will experience learning content as well as assessment content. Assessments are designed to measure the effectiveness of learning content and help assess knowledge gained. It is possible that certain assessment questions do not accurately represent the concepts contained within learning content and this may be indicated by a majority of learners getting the question wrong. It is also possible that the question accurately represents the learning content but is very difficult. The following algorithm will identify these types of questions but will not be able to deduce why learners answer them incorrectly.

### 14.1 Ideal Statements

In order to accurately determine which assessment questions are the most difficult, there are a few requirements of the data produced by a LRP. They are as follows:

- statements describing a learner answering a question must report if the learner got the question correct or incorrect via *\$.result.success*
- if it is possible to get partial credit on a question, the amount of credit should be reported within the statement
  - the credit earned by the learner should be reported within *\$.result.score.raw*
  - the minimum and maximum possible credit amount should be reported within *\$.result.score.min* and *\$.result.score.max* respectively
- If it is possible to get partial credit on a question, it must still be reported if the learner reached the threshold of success via *\$.result.success*
- Statements describing a learner answering a question should contain activities of the type *cmi.interaction*
- activities must be uniquely and consistently identified across all statements
- Statements describing a learner answering a question should<sup>10</sup> use the verb *http://adlnet.gov/expapi/verbs/answered*

---

<sup>10</sup> it is possible to use another verb *iri* but if another is used, that will need to be accounted for in data retrieval

## 14.2 Input Data Retrieval

How to query an LRS via a GET request to the Statements Resource via curl. The following section contains the appropriate parameters with example values as well as the curl command necessary for making the request.<sup>111213</sup>

```
Verb = "verb=http://adlnet.gov/expapi/verbs/answered"

Since = "since=2018-07-20T12:08:47Z"

Until = "until=2018-07-21T12:08:47Z"

Base = "https://example.endpoint/statements?"

endpoint = Base + Verb + "&" + Since + "&" + Until

Auth = Hash generated from basic auth

S = curl -X GET -H "Authorization: Auth"
-H "Content-Type: application/json"
-H "X-Experience-API-Version: 1.0.3"
Endpoint
```

## 14.3 Statement Parameters to Utilize

The statement parameter locations here are written in [JSONPath](#). This notation is also compatible with the xAPI Z notation due to the defined hierarchy of components. Within the Z specifications, a variable name will be used instead of the \$

- *\$.result.success*
- *\$.object.id*

## 14.4 2018 Pilot TLA Statement Problems

The initial pilot test data supports this algorithm. Given that the official 2018 pilot test is scheduled to take place on July 27th, 2018, this section may require updates pending future data review.

## 14.5 Summary

1. Query an LRS via a [GET](#) request to the statements endpoint using the parameters verb, since and until

---

<sup>11</sup> See footnote 1.

<sup>12</sup> See footnote 2.

<sup>13</sup> See footnote 3.

2. Filter the results to the set of statements where:

- $\$.result.success$  is false

3. process the filtered data

- group by  $\$.object.id$
- determine the count of each group
- create a collection of pairs =  $[\$.object.id, \#]$

## 14.6 Formal Specification

### 14.6.1 Basic Types

$INCORRECT ::= \{false\}$

### 14.6.2 System State

$MostDifficultAssessmentQuestions$
$Statements$
$S_{all} : \mathbb{F}_1$
$S_{incorrect}, S_{grouped}, S_{processed} : \mathbb{F}$
$S_{all} = statements$
$S_{incorrect} \subseteq S_{all}$
$S_{grouped} = \{groups : seq_1 statement\}$
$S_{processed} = \{pair : (id, \mathbb{N})\}$

- The set  $S_{all}$  is a non-empty, finite set and is the component *statements*
- The sets  $S_{incorrect}$ ,  $S_{grouped}$  and  $S_{processed}$  are all finite sets
- the set  $S_{incorrect}$  is a subset of  $S_{all}$  which may contain every value within  $S_{all}$
- the set  $S_{grouped}$  is a finite set of objects *groups* which are non-empty, finite sequences of the component *statement*
- the set  $S_{processed}$  is a finite set of pairs where each contains the component *id* and a natural number

### 14.6.3 Initial System State

$InitMostDifficultAssessmentQuestions$
$MostDifficultAssessmentQuestions$
$S_{all} \neq \emptyset$
$S_{incorrect} = \emptyset$
$S_{grouped} = \emptyset$
$S_{processed} = \emptyset$



- The set  $S_{all}$  is a non-empty set
- The sets  $S_{incorrect}$  ,  $S_{grouped}$  and  $S_{processed}$  are all initially empty

#### 14.6.4 Filter for Incorrect

$Incorrect$ $Statement$ $incorrect : STATEMENT \rightarrow \mathbb{F}$ $s? : STATEMENT$ $s! : \mathbb{F}$
$s? = statement$ $s! = incorrect(s?)$ $incorrect(s?) = \text{if } s?.result.success : INCORRECT$ $\quad \text{then } s! = s?$ $\quad \text{else } s! = \emptyset$

- the schema  $Incorrect$  introduces the function  $incorrect$  which takes in the variable  $s?$  and returns the variable  $s!$
- the variable  $s?$  is the component  $statement$
- $s!$  is equal to  $s?$  if  $s?.result.success$  is of the type  $INCORRECT$  otherwise  $s!$  is an empty set

$FilterForIncorrect$ $\Delta MostDifficultAssessmentQuestions$ $Incorrect$ $incorrects : \mathbb{F}$
$incorrects \subseteq S_{all}$ $incorrects' = \{s : STATEMENT \mid incorrect(s) \neq \emptyset\}$ $S'_{incorrect} = S_{incorrect} \cup incorrects'$

- the set  $incorrects$  is a subset of  $S_{all}$  which may contain every value within  $S_{all}$
- The set  $incorrects'$  contains elements  $s$  of type  $STATEMENT$  where  $incorrect(s)$  is not an empty set
- The updated set  $S'_{incorrect}$  is the union of the previous state of  $S_{incorrect}$  and  $incorrects'$

#### 14.6.5 Processes Results

<i>GroupByActivityId</i>	
<i>Statements</i>	
$g? : \mathbb{F}$	
$g! : \mathbb{F}$	
$group : \mathbb{F} \rightarrow \mathbb{F}$	
$g? = statements \Rightarrow \{g : statement\}$	
$g! = group(g?)$	
$g! = \{groups : seq_1 statement \mid$	
<b>let</b> $seq_1 statement_i..statement_j == seq_1 s_i..s_j \bullet$	
$\forall s_n : s_i..s_j \bullet i \leq n \leq j \bullet s_i.object.id = s_j.object.id = s_n.object.id\}$	

- The schema *GroupByActivityId* introduces the function *group* which has the input of *g?* and the output of *g!*
- The input variable *g?* is the component *statements* which implies its a set of objects *g* which are each a *statement*
- the output variable *g!* is a set of objects *groups* which are each a non-empty, finite sequence of *statement* where each member of the sequence  $s_i..s_j$  has the same \$.object.id

<i>CountPerGroup</i>	
<i>Statement</i>	
$c? : seq_1 statement$	
$c! : \mathbb{N}$	
$count : seq_1 statement \rightarrow \mathbb{N}$	
$c! = count(c?)$	
$c! \geq 1$	
$count(c?) = \forall c_n? : \langle c?_i .. c?_j \rangle \bullet i \leq n \leq j \wedge i = 0 \bullet$	
$\exists_1 c! : \mathbb{N} \bullet \text{if } n = i \text{ then } c! = n + 1 \text{ else } c! = j + 1$	

- The schema *CountPerGroup* introduces the function *count* which has the input of *c?* and the output of *c!*
- The input variable *c?* is a non-empty, finite sequence in which each element is a *statement*
- The function *count* reads: for all elements  $c?_n$  within the sequence  $\langle c?_i .. c?_j \rangle$ , such that *n* is greater than or equal to *i* and less than or equal to *j*, *i* is equal to zero and there exists a number *c!* which is equal to *n* + 1 (when  $n = i \Rightarrow n = 0$ ) or equal to *n*

---

*AggregateQuestionStatements*  
 $\Delta$ *MostDifficultAssessmentQuestions*  
*FilterForIncorrect*  
*GroupByActivityId*  
*CountPerGroup*  
*grouped, processed* :  $\mathbb{F}$

---

$grouped = \emptyset$   
 $grouped' = group(S_{incorrect})$   
 $S'_{grouped} = S_{grouped} \cup grouped'$   
 $processed \subseteq S'_{grouped}$   
 $processed' = \{p : (id, \mathbb{N}) \mid$   
 $\quad \text{let } \{(processed_i) .. (processed_j)\} == \{g_i .. g_j\} \bullet$   
 $\quad i \leq n \leq j \bullet \forall g_n : g_i .. g_j \bullet \exists p_n : p_i .. p_j \bullet$   
 $\quad first\ p_n = head\ g_n.object.id \wedge second\ p_n = count(g_n)$   
 $\left. S'_{processed} = S_{processed} \cup processed' \right\}$

---

- The schema *AggregateQuestionStatements* introduces the variables *grouped* and *processed*
- *grouped* starts as an empty set but then becomes *grouped'* which is the output of applying the function *group* to the set of statements *S<sub>incorrect</sub>* created by the operation *FilterForIncorrect*
- *grouped'* is a set of sequences. The elements of those sequences are statements which all have the same *statement.object.id*
- The set *S<sub>grouped</sub>* is updated to the set *S'<sub>grouped</sub>* which is the union of *S<sub>grouped</sub>* and *grouped'*
- the variable *processed* is a subset of *S'<sub>grouped</sub>* which can contain every value within *S'<sub>grouped</sub>*
- the variable *processed* is updated to be the variable *processed'* which is a set of objects *p* which are ordered pairs of the component *id* and a natural number. *p* is defined as:
  - for all sequences *g<sub>i</sub> .. g<sub>j</sub>* within the set *processed*, there exists an ordered pair *p<sub>n</sub>* such that:
    - \* the first element of *p<sub>n</sub>* is equal to the *object.id* of the first statement within the sequence *g<sub>n</sub>*.
    - \* The second element of *p<sub>n</sub>* is equal to the value returned when *g<sub>n</sub>* is passed to the function *count*.
- The set *S'<sub>processed</sub>* is the union of the sets *S<sub>processed</sub>* and *processed'*

### 14.6.6 Sequence of Operations

$ProcessedQuestions \hat{=} FilterForIncorrect \circ AggregateQuestionStatements$

- The schema  $ProcessedQuestions$  is the sequential composition of operation schemas  $FilterForIncorrect$  and  $AggregateQuestionStatements$
- $FilterForIncorrect$  happens before  $AggregateQuestionStatements$

### 14.6.7 Return

$ReturnAggregate$ $\exists MostDifficultAssessmentQuestions$ $ProcessedQuestions$ $S_{processed}! : \mathbb{F}$	
$S_{processed}! = S_{processed}$	

- The returned variable  $S_{processed}!$  is equal to the current state of variable  $S_{processed}$  after the operations  $FilterForIncorrect$  and  $AggregateQuestionStatements$

## 14.7 Pseudocode

---

### Algorithm 3: Most Difficult Assessment Questions

---

```

Input:  $S_{all}, displayN$ 
Result:  $display''$ 
 $context = \{\}$ ;
 $display = []$ ;
while  $S_{all} \neq \emptyset$  do
    foreach  $s \in S_{all}$  do
        if  $s.result.success = INCORRECT$  then
            do
                 $S'_{incorrect} \leftarrow s \cup S_{incorrect}$ ;
                 $S'_{all} \leftarrow S_{all} \setminus s$ ;
                recur  $S'_{all}, S'_{incorrect}$ ;
            else
                do
                     $S'_{all} \leftarrow S_{all} \setminus s$ ;
                    recur  $S'_{all}$ 
                end
            end
        end
    end
    while  $S'_{incorrect} \neq \emptyset$  do
        foreach  $si \in S'_{incorrect}$  do
             $id \leftarrow si.object.id$ ;
            if  $id \notin context$  then
                do
                     $count = 1$ ;
                     $context' \leftarrow \{id : count\}$ ;
                     $S'_{incorrect} \leftarrow S_{incorrect} \setminus si$ ;
                    recur  $context', S'_{incorrect}$ ;
                else
                    do
                         $count' \leftarrow inc(context.id)$ ;
                         $context' \leftarrow \{id : count'\}$ ;
                         $S'_{incorrect} \leftarrow S_{incorrect} \setminus si$ ;
                        recur  $context', S'_{incorrect}$ ;
                    end
                end
            end
        end
    end
    foreach  $id \in context'$  do
         $IdToCount \leftarrow [id, context.id]$ ;
         $display' \leftarrow display \cap IdToCount$ ;
        recur  $display'$ 
    end
return  $display'' \leftarrow take(sortBySubArray(display'), displayN)$ 

```

---

- The Z schemas are used within this pseudocode
- The return value display is an array of length display-n, where each element of display is an array of  $[statement.object.id, \#]$  where  $\#$  representing the number of times  $statement.object.id$  appeared within  $S'_{incorrect}$

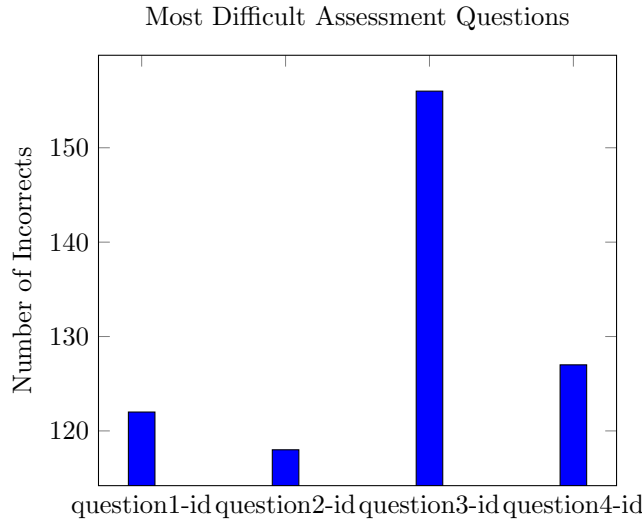
## 14.8 JSON Schema

```
{ "type": "array",
  "items": { "type": "array",
    "items": [ { "type": "string" }, { "type": "number" } ] } }
```

## 14.9 Visualization Description

The **Most Difficult Assessment Questions** visualization will be a bar chart where the domain consists of  $statement.object.id$  and the range is a number greater than or equal to 1. Every subarray within the array display will be a grouping within the bar chart. The pseudocode specifies an input parameter display-n which controls the length of the array display and therefore the number of groups contained within the visualization.

## 14.10 Visualization prototype



## 14.11 Prototype Improvement Suggestions

Additional features may be implemented on top of this base specification but they would require adding additional values to each subarray returned by the algorithm. These additional values can be retrieved via (1) performing metadata lookup within or independently of the algorithm (2) by utilizing additional xAPI

statement parameters and/or (3) by performing additional computations. The following examples assume the metadata is contained within each statement available to the algorithm.

- Use the name of the activity for the x-axis label instead of its id.
  - *\$.object.definition.name*
  - grouping of statements should still happen by *\$.object.id* to ensure an accurate count
- a tooltip containing contextual information about the question such as:
  - The question text
    - \* *\$.object.definition.description*
  - Interaction Type
    - \* *\$.object.definition* which contains interaction properties
  - Answer choices
    - \* *\$.object.definition* which contains interaction properties
  - Correct answer
    - \* *\$.object.definition* which contains interaction properties
  - Most popular incorrect answer
    - \* This would require an extra step of processing and all statements would need to utilize interaction properties within *\$.object.definition*
  - average partial credit earned (if applicable)
    - \* *\$.result.score.scaled*
    - \* The one potential issue with using scaled score is the calculation of scaled is not strictly defined by the xAPI specification but is instead up to the authors of the LRP. This results in the inability to reliably compare scaled scores across LRPs.
    - \* if *\$.result.score.raw* , *\$.result.score.min* and *\$.result.score.max* are reported for all questions, it becomes possible to reliably compare scores across questions and LRPs.
  - average number of re-attempts
    - \* this would require additional steps of processing so that *\$.actor* is considered as well
    - \* due to the problem of actor unification, ie the same person being identified differently across statements, this metric may not be accurate.
  - average time spent on the question
    - \* *\$.result.duration*
    - \* this would require additional steps of processing to extract the duration and average it.

- a tooltip containing contextual information about the course and/or assessment the question was within
  - the instructor for the course
    - \* *\$.context.instructor*
  - competency associated with the question and/or course
    - \* *\$.context.contextActivities*
  - metadata about the learning content associated with the question such as average time spent engaging with associated content before attempting the question.
  - this would require additional steps of processing to retrieve metadata about the content and its usage.
    - \* *\$.context.contextActivities*



## 15 How Often are Recommendations Followed

As learners engage in activities supported by a learning ecosystem, they will build up a history of learning experiences. When the digital resources of that learning ecosystem adhere to a framework dedicated to supporting and understanding the learner, such as the Total Learning Architecture (TLA), it becomes possible to retell their learning story through data and data visualization. One important aspect of that story is the recommendations provided to the learner and whether or not the learner follows those recommendations.

### 15.1 Ideal Statements

In order to accurately determine if a learner is following recommendations, there are a few requirements of the data produced by a LRP and the recommender itself. They are as follows:

- Every time the recommender makes a recommendation, a statement should be produced which uses the verb `https://w3id.org/xapi/dod-isd/verbs/recommended`<sup>14</sup> and has the recommended piece of content as the object.
  - the content should be uniquely and consistently identified across all statements.
- When a learner launches recommended content, the resulting launched statement should use the verb `http://adlnet.gov/expapi/verbs/launched`<sup>15</sup> and contain a reference to the recommended content statement within `$.context.statement`
  - Launching of content should use the above IRI regardless of why the content was launched
  - If it not possible to reference the exact recommended content statement, the launch statement should have some indication that it was the result of a recommendation.<sup>16</sup>

---

<sup>14</sup> See footnote 4

<sup>15</sup> See footnote 4

<sup>16</sup> It is possible to determine if recommendations are followed (with some level of error) without this explicit linking of launched to recommended but this severely complicates the algorithm. In this case, in order to optimize for accuracy, the algorithm would need to consider the actor and their general activity within a session, the object of both launched and recommended statements generated within the session, the time lapse between recommendations and launches with a predefined lapse value which determines if a launch was close enough in time to a recommendation to be considered a result of the recommendation. An additional constraint on the above case is the recommendation statements should contain a reference to the person receiving the recommendation, otherwise determining the 1:1 relationships between recommendations and launches requires additional complexity and will still not be 100% accurate due to the reliance on the time lapse value.

## 15.2 Input Data Retrieval

How to query an LRS via a GET request to the Statements Resource via curl.<sup>171819</sup>

```
R = "verb=https://w3id.org/xapi/dod-isd/verbs/recommended"
L = "verb=http://adlnet.gov/expapi/verbs/launched"

Since = "since=2018-07-20T12:08:47Z"
Until = "until=2018-07-21T12:08:47Z"
Base = "https://example.endpoint/statements?"

endpoint1 = Base + R + "&" + Since + "&" + Until
endpoint2 = Base + L + "&" + Since + "&" + Until

Auth = Hash generated from basic auth

SR = curl -X GET -H "Authorization: Auth"
      -H "Content-Type: application/json"
      -H "X-Experience-API-Version: 1.0.3"
      endpoint1
SL = curl -X GET -H "Authorization: Auth"
      -H "Content-Type: application/json"
      -H "X-Experience-API-Version: 1.0.3"
      endpoint2

S = SR + SL
```

## 15.3 Statement Parameters to Utilize

The statement parameter locations here are written in [JSONPath](#). This notation is also compatible with the xAPI Z notation due to the defined hierarchy of components. Within the Z specifications, a variable name will be used instead of the \$

- *\$.verb.id*
- *\$.context.statement*

## 15.4 2018 Pilot TLA Statement Problems

At the time of writing this document, launched statements do not include a statement reference or any indication of a connection between recommendations and launches. The authors of this document do not have access to the LRS

<sup>17</sup> footnote 1 applies to both S1 and S2.

<sup>18</sup> See footnote 2.

<sup>19</sup> See footnote 3.

containing the recommended statements and thus can not draw any conclusions about any issues which may be present within those statements or any aspects of those statements which may correlate them to launch statements. The following algorithm is going to assume that the input set of statements follow the guidelines outlined in section 5.1 as the additional algorithmic considerations brought on by non ideal statements, as specified within footnote 16, result in an algorithm which is not optimal for near real time visualizations.

## 15.5 Summary

1. Query an LRS via a GET request to the statements endpoint using the parameters verb, since and until to gather all statements with the verb *http : //adlnet.gov/expapi/verbs/launched*.
2. Query an LRS via a GET request to the statements endpoint using the parameters verb, since and until to gather all statements with the verb *https : //w3id.org/xapi/dod - isd/verbs/recommended*.<sup>20</sup>
3. Group all collections of statements by a *TIMEUNIT*
4. separate the collection of grouped launched statements into a collection of those which were the result of a recommendation and those which were not.
5. Take the count of all groups of statements
  - Recommended statements per *TIMEUNIT*
  - Launches due to recommendations per *TIMEUNIT*
  - Launches not due to recommendations per *TIMEUNIT*
6. Calculate summary statistics for the overall time range and per *TIMEUNIT*
  - Divide launches due to recommendations by the total number of launches to determine the percentage of launches due to recommendations
  - Divide launches due to recommendations by the total number of recommendations to determine the percentage of recommendations which are followed.

---

<sup>20</sup> If since and until are specified, they should be the same in both requests.

## 15.6 Formal Specification

### 15.6.1 System State

<i>FollowedRecommendations</i>	_____
<i>Statements</i>	
<i>CountPerGroup</i>	
$S_{recommended}, S_{launched} : \mathbb{F}_1$	
$ordered_L, ordered_R, grouped_{launched}, grouped_{recommended},$	
$onlyRecommended, cPerGroup_{launched}, cPerGroup_{recommended},$	
$cPerGroup_{followed}, combined : seq$	
$t_{start}, N_{launched}, N_{recommended}, N_{followed}, P_{followed}, P_{dueto} : \mathbb{N}$	
$tr_{start}, tr_{end} : \mathbb{F}$	
$unit? : TIMEUNIT$	
$S_{recommended} = statements$	
$S_{launched} = statements$	
$combined = \langle (tr_{start}, tr_{end}, N_{launched}, N_{recommended}, N_{followed}, P_{followed}, P_{dueto}) \rangle$	
$count(grouped_{launched}) = count(grouped_{recommended})$	
$count(onlyRecommended) = count(grouped_{launched}) \Rightarrow$	
$count(onlyRecommended) = count(grouped_{recommended})$	
$count(cPerGroup_{launched}) = count(cPerGroup_{followed}) = count(cPerGroup_{recommended})$	

- $S_{recommended}, S_{launched}$  are both non-empty, finite sets.
  - $S_{recommended}$  and  $S_{launched}$  contain the results of querying an LRS for recommended and launched statements respectively.
- $ordered_L, ordered_R, grouped_{launched}, grouped_{recommended}, onlyRecommended, cPerGroup_{launched}, cPerGroup_{recommended}, cPerGroup_{followed}$  and  $combined$  are all finite sequences.
  - $ordered_L$  and  $ordered_R$  are the sequences of statements within  $S_{launched}$  and  $S_{recommended}$  respectively and sorted by timestamp.
  - $grouped_{launched}$  is the result of grouping the statements within  $ordered_L$  by  $unit?$ .
  - $grouped_{recommended}$  is the result of grouping the statements within  $ordered_R$  by  $unit?$ .
  - $onlyRecommended$  is the result of filtering the statements within the sequence  $grouped_{launched}$  to only include statements where  $statement.context.statement$  is present
  - $cPerGroup_{launched}, cPerGroup_{recommended}, cPerGroup_{followed}$  are all sequences of numbers which represent the count within each subsequence of  $grouped_{launched}, grouped_{recommended}$  and  $onlyRecommended$  respectively.

- *combined* is a sequence of ordered pairs where each pair consists of  $tr_{start}$ ,  $tr_{end}$ ,  $N_{launched}$ ,  $N_{recommended}$ ,  $N_{followed}$ ,  $P_{followed}$  and  $P_{dueto}$
- $t_{start}$ ,  $N_{launched}$ ,  $N_{recommended}$ ,  $N_{followed}$ ,  $P_{followed}$ ,  $P_{dueto}$  are all natural numbers
- $tr_{start}$ ,  $tr_{end}$  are both timestamps which represent the the start and end of the time range for each a group of statements.
- *unit?* is an input representing a time interval, ie day vs month vs hour.
- all sequences are the same length so that each subsequence represents the same time grouping. In other words, indexes are comparable across sequences.

### 15.6.2 Initial System State

<i>InitFollowedRecommendations</i>	_____
<i>FollowedRecommendations</i>	_____
$S_{recommended} \neq \emptyset$	
$S_{launched} \neq \emptyset$	
$unit? = \{day\}$	
$ordered_L = \langle \rangle$	
$ordered_R = \langle \rangle$	
$grouped_{launched} = \langle \rangle$	
$grouped_{recommended} = \langle \rangle$	
$onlyRecommended = \langle \rangle$	
$cPerGroup_{launched} = \langle \rangle$	
$cPerGroup_{recommended} = \langle \rangle$	
$cPerGroup_{followed} = \langle \rangle$	
$combined = \langle \rangle$	
$t_{start} = 0$	
$N_{launched} = 0$	
$N_{recommended} = 0$	
$N_{followed} = 0$	
$P_{followed} = 0$	
$P_{dueto} = 0$	

- $S_{recommended}$  and  $S_{launched}$  are initially not empty sets
- all sequences are initially empty
- all numbers are initially zero
- the default *TIMEUNIT* is set to day

### 15.6.3 Group by Timestamp

$SortByTimestamp$ $Statement$ $IsoToUnix$ $orderByTimestamp : \mathbb{F}_1 \rightarrow seq_1$ $o? : \mathbb{F}_1$ $o! : seq_1 statement$
$o? = \{o : statement\}$ $o! = orderByTimestamp(o?)$ $o! = \langle o_i..o_j \rangle \bullet \forall o_n : o_i..o_j \bullet o_n : STATEMENT \wedge i \leq n \leq j \bullet$ $convert(o_i.timestamp) \leq convert(o_n.timestamp) \leq convert(o_j.timestamp)$

- The schema *SortByTimestamp* introduces the function *orderByTimestamp* which takes in a non-empty, finite set and returns a non-empty, finite sequence.
- *orderByTimestamp* is a sequence of statements ordered from earliest to latest.

$WithinRange$ $withinRange : (\mathbb{N}, \mathbb{N}, \mathbb{N}, TIMEUNIT) \rightarrow \mathbb{F}_1 \#1$ $in?, start?, state? : \mathbb{N}$ $unit? : TIMEUNIT$ $out! : \{TRUE\} \vee \{FALSE\}$
$unit? = \{second\} \Rightarrow 1 \vee \{minute\} \Rightarrow 60 \vee \{hour\} \Rightarrow 3600 \vee$ $\{day\} \Rightarrow 86400 \vee \{week\} \Rightarrow 604800 \vee$ $\{month\} \Rightarrow 2629743 \vee \{year\} \Rightarrow 31556926$ $out! = withinRange(in?, start?, state?, unit?)$ $withinRange(in?, start?, state?, unit?) =$ $\quad \text{if } in? \leq start? + ((state? + 1) * unit?)$ $\quad \quad \text{then } out! = \{TRUE\}$ $\quad \quad \text{else } out! = \{FALSE\}$

- The schema *WithinRange* introduces the function *withinRange* which takes in three numbers and a *TIMEUNIT* and returns either  $\{TRUE\}$  or  $\{FALSE\}$
- *withinRange* checks to see if *in?* is less than or equal to a start time *start?* plus the result of multiplying the numeric conversion for *unit?* by the *state?*.
- *state?* represents the current group, ie. day 1 vs day 2 vs day 3. The +1 is to account for array indexes starting at 0.

<i>GroupByTimeUnit</i>	
<i>Statement</i>	
<i>IsoToUnix</i>	
<i>WithinRange</i>	
$groupByTimeUnit : (\text{seq}_1, \mathbb{N}, \text{TIMEUNIT}) \rightarrow \text{seq}_1$	
$g?, g! : \text{seq}_1$	
$t_{start?} : \mathbb{N}$	
$ \begin{aligned} &g? = \langle g?_i .. g?_j \rangle \bullet \forall g?_n : g?_i .. g?_j \bullet i \leq n \leq j \bullet g?_n = \text{statement} \wedge \\ &\quad \text{convert}(g?_i.\text{timestamp}) \leq \text{convert}(g?_n.\text{timestamp}) \leq \text{convert}(g?_j.\text{timestamp}) \\ &g! = \text{groupByTimeUnit}(g?, t_{start?}, \text{state?}, \text{unit?}) \\ &g! = \langle g : \text{seq} \mid \forall g?_n : g?_i .. g?_j \bullet \exists_1 \langle g_r \rangle : \langle g_q \rangle .. \langle g_s \rangle \bullet q \leq r \leq s \wedge r = \text{state?} \bullet \\ &\quad \text{if } \text{withinRange}(\text{convert}(g?_n.\text{timestamp}), t_{start?}, r, \text{unit?}) = \{TRUE\} \\ &\quad \quad \text{then } g? \upharpoonright g?_n \wedge g?_n \text{ in } \langle g_r \rangle \Rightarrow \langle g_r \rangle = \langle g_{ri} .. g_{rn} .. g_{rj} \rangle \bullet ri \leq rn \leq rj \bullet g_{rn} = g?_n \\ &\quad \quad \text{else if } \forall g_n? : g_i? .. g_j? \bullet \text{withinRange}(g_n?, t_{start?}, r, \text{unit?}) = \{FALSE\} \\ &\quad \quad \text{then } \text{groupByTimeUnit}((g? \upharpoonright \langle g_r \rangle), t_{start?}(r+1), \text{unit?}) \bullet \langle g_r \rangle = \langle \rangle \vee \neq \langle \rangle \end{aligned} $	

- The schema *GroupByTimeUnit* introduces the function *groupByTimeUnit* which takes as arguments a non-empty, finite sequence, a natural number and a *TIMEUNIT* and outputs a non-empty, finite sequence of sequences.
- For every statement within the input sequence, *groupByTimeUnit* checks to see if the timestamp of that statement is within the range of  $t_{start}$  and  $unit?$ . If it is, that statement is removed from the input sequence  $g?$  and added to the current subsequence  $\langle g_r \rangle$ . If none of the remaining statements within the input sequence are within the range of  $t_{start}$  and  $unit?$ , then the variable  $state?$  is incremented, the current subsequence  $\langle g_r \rangle$  is either a collection of matched statements or is an empty sequence and the search for remaining subsequences  $\langle g_{r+state?} \rangle$  continues.
- because the input sequence  $g?$  is ordered chronologically, this implies that once a statement does not fit into a range, the rest of the statements remaining in the input sequence will not fit into that range and  $state?$  must be incremented to generate a new subsequence  $\langle g_{r+state?} \rangle$  so that the remaining statements can be grouped.

#### 15.6.4 Processes Results

<i>OrderStatements</i>	_____
$\Delta$ <i>FollowedRecommendations</i>	
<i>SortByTimestamp</i>	
<hr/>	
$ordered'_L = orderByTimestamp(S_{launched})$	
$ordered'_R = orderByTimestamp(S_{recommended})$	
$t'_{start} = convert((head\ ordered'_L).timestamp)$	
<hr/>	

- The schema *OrderStatements* updates the system state defined by the schema *FollowedRecommendations*.
- $ordered'_L$  is the result of ordering the statements contained within the set  $S_{launched}$  chronologically.
- $ordered'_R$  is the result of ordering the statements contained within the set  $S_{recommended}$  chronologically.
- $t'_{start}$  is the timestamp from the first statement within  $ordered'_L$  converted to unix time.

<i>GroupByTime</i>	_____
$\Delta$ <i>FollowedRecommendations</i>	
<i>GroupByTimeUnit</i>	
<hr/>	
$grouped'_{launched} = groupByTimeUnit(ordered'_L, t'_{start}, 0, unit?)$	
$grouped'_{recommended} = groupByTimeUnit(ordered'_R, t'_{start}, 0, unit?)$	
<hr/>	

- The schema *GroupByTime* updates the state defined by the schema *FollowedRecommendations*.
- $grouped'_{launched}$  is the result of passing  $ordered'_L$ ,  $t'_{start}$ , 0 and *unit?* to the function *groupByTimeUnit*.
- $grouped'_{recommended}$  is the result of passing  $ordered'_R$ ,  $t'_{start}$ , 0 and *unit?* to the function *groupByTimeUnit*.

<i>OnlyRecommendedLaunches</i>	_____
$\Delta$ <i>FollowedRecommendations</i>	
<hr/>	
$onlyRecommended' = \langle o : seq \mid \mathbf{let} \ grouped'_{launched} == gl \Rightarrow$	
$\langle \langle gl_i \rangle .. \langle gl_j \rangle \rangle \Rightarrow \langle \langle gl_{ii} .. gl_{ij} \rangle .. \langle gl_{ji} .. gl_{jj} \rangle \rangle \bullet$	
$\forall \langle gl_n \rangle : \langle gl_i \rangle .. \langle gl_j \rangle \bullet \exists_1 \langle o_n \rangle : \langle o_i \rangle .. \langle o_j \rangle \Rightarrow \langle \langle o_{ii} .. o_{ij} \rangle .. \langle o_{ji} .. o_{jj} \rangle \rangle \bullet$	
$((\forall o_{in} : o_{ii} .. o_{ij} \bullet o_{in}.context.statement \neq \emptyset \wedge o_{in} \text{ in } gl_i) \wedge$	
$(\forall o_{jn} : o_{ji} .. o_{jj} \bullet o_{jn}.context.statement \neq \emptyset \wedge o_{jn} \text{ in } gl_j)) \vee$	
$\langle o_n \rangle = \langle \rangle \rangle$	
<hr/>	



- The schema *OnlyRecommendedLaunches* updates the state defined by the schema *FollowedRecommendations*.
- *onlyRecommended'* is the sequence of objects  $o$  where  $o$  is a sequence consisting of statements (or no statements) from the corresponding sequences within  $grouped'_{launched}$  where  $statement.context.statement$  exists.
- *onlyRecommended'* maintains the same number and ordering of time groups (subsequences) as  $grouped'_{launched}$  and  $grouped'_{recommended}$ .

<i>GetCounts</i>	
$\Delta$ <i>FollowedRecommendations</i>	
<i>CountPerGroup</i>	
$cPerGroup'_{launched} = \langle c : \mathbb{N} \mid \text{let } grouped'_{launched} == gl \Rightarrow \langle \langle gl_i \rangle .. \langle gl_j \rangle \rangle \bullet$	
$\quad \forall \langle gl_n \rangle : \langle gl_i \rangle .. \langle gl_j \rangle \bullet \exists_1 c_n : \mathbb{N} \bullet$	
$\quad \text{if } gl_n = \langle \rangle$	
$\quad \quad \text{then } c_n = 0$	
$\quad \quad \text{else } c_n = count(\langle gl_n \rangle)$	
$cPerGroup'_{recommended} = \langle c : \mathbb{N} \mid \text{let } grouped'_{recommended} == gr \Rightarrow \langle \langle gr_i \rangle .. \langle gr_j \rangle \rangle \bullet$	
$\quad \forall \langle gr_n \rangle : \langle gr_i \rangle .. \langle gr_j \rangle \bullet \exists_1 c_n : \mathbb{N} \bullet$	
$\quad \text{if } gr_n = \langle \rangle$	
$\quad \quad \text{then } c_n = 0$	
$\quad \quad \text{else } c_n = count(\langle gr_n \rangle)$	
$cPerGroup'_{followed} = \langle c : \mathbb{N} \mid \text{let } onlyRecommended' == or \Rightarrow \langle \langle or_i \rangle .. \langle or_j \rangle \rangle \bullet$	
$\quad \forall \langle or_n \rangle : \langle or_i \rangle .. \langle or_j \rangle \bullet \exists_1 c_n : \mathbb{N} \bullet$	
$\quad \text{if } or_n = \langle \rangle$	
$\quad \quad \text{then } c_n = 0$	
$\quad \quad \text{else } c_n = count(\langle or_n \rangle)$	

- The schema *GetCounts* updates the state defined by the schema *FollowedRecommendations*.
- $cPerGroup'_{launched}$  is a sequence of numbers  $c$  where each  $c$  is either 0 or the result of passing the current subsequence of  $grouped'_{launched}$  ( $gl_n$ ) to the function *count*.
- $cPerGroup'_{recommended}$  is a sequence of numbers  $c$  where each  $c$  is either 0 or the result of passing the current subsequence of  $grouped'_{recommended}$  ( $gr_n$ ) to the function *count*.
- $cPerGroup'_{followed}$  is a sequence of numbers  $c$  where each  $c$  is either 0 or the result of passing the current subsequence of *onlyRecommended'* ( $or_n$ ) to the function *count*.

*CombineSequences*

$\Delta$ *FollowedRecommendations*

```

combined' = ⟨c : (tr'_start, tr'_end, N'_launched, N'_recommended, N'_followed, P'_followed, P'_dueto) |
  let grouped'_launched == gl ⇒ ⟨⟨gl_i⟩..⟨gl_n⟩..⟨gl_j⟩⟩
  cPerGroup'_launched == cl ⇒ ⟨cl_i..cl_n..cl_j⟩
  cPerGroup'_recommended == cr ⇒ ⟨cr_i..cr_n..cr_j⟩
  cPerGroup'_followed == cf ⇒ ⟨cf_i..cf_n..cf_j⟩
  • ∀⟨gl_n⟩ : ⟨gl_i⟩..⟨gl_j⟩ • i ≤ n ≤ j •
  ∃_1 c_n : (tr_startn, tr_endn, N_launchedn, N_recommendedn, N_followedn, P_followedn, P_dueton) •
  tr_startn = (head gl_n).timestamp
  tr_endn = (last gl_n).timestamp
  N_launchedn = cl_n
  N_recommendedn = cr_n
  N_followedn = cf_n
  P_followedn = cf_n ÷ cr_n
  P_dueton = cf_n ÷ cl_n⟩

```

- The schema *CombineSequences* changes the state defined by the schema *FollowedRecommendations*.
- *combined'* is a sequence of objects *c* where each *c* is an ordered pair of  $tr'_start, tr'_end, N'_launched, N'_recommended, N'_followed, P'_followed, P'_dueto$ .
- for each  $c_n$ :
  - $tr'_start \rightsquigarrow tr\_startn$  which is equal to the timestamp for the first statement within  $gl_n$
  - $tr'_end \rightsquigarrow tr\_endn$  which is equal to the timestamp for the last statement within  $gl_n$ .
  - $N'_launched \rightsquigarrow N\_launchedn$  which is equal to the current count of launched statements within the nth time grouping aka  $cl_n$ .
  - $N'_recommended \rightsquigarrow N\_recommendedn$  which is equal to the current count of recommended statements within the nth time grouping aka  $cr_n$ .
  - $N'_followed \rightsquigarrow N\_followedn$  which is equal to the current count of recommended statements within the nth time grouping aka  $cf_n$ .
  - $P'_followed \rightsquigarrow P\_followedn$  which is equal to the result of dividing  $cf_n$  by  $cr_n$ .
  - $P'_dueto \rightsquigarrow P\_dueton$  which is equal to the result of dividing  $cf_n$  by  $cl_n$ .

### 15.6.5 Sequence of Operations

*ProcessFollowedRecommendations*  $\hat{=}$   
*OrderStatements* ; *GroupByTime* ; *OnlyRecommendedLaunches* ;  
*GetCounts* ; *CombineSequences*

- The schema *ProcessFollowedRecommendations* defines the order of operations for the steps within the *FollowedRecommendations* algorithm.

#### 15.6.6 Return

<i>ReturnFollowedRecommendations</i>	_____
$\Xi$ <i>FollowedRecommendations</i>	
<i>ProcessFollowedRecommendations</i>	
<i>combined!</i> : seq	
<i>combined!</i> = <i>combined'</i>	_____

- The schema *ReturnFollowedRecommendations* describes the return value of the system defined by the schema *FollowedRecommendations*
- The return value *combined!* is the variable *combined'* defined within the schema *CombineSequences*

## 15.7 Pseudocode

---

**Algorithm 4:** Followed Recommendations

---

**Input:**  $S_{recommended}$ ,  $S_{launched}$  *timeUnit*  
**Result:** *combined'*  
 $ordered'_L \leftarrow orderByTimestamp(S_{launched});$   
 $ordered'_R \leftarrow orderByTimestamp(S_{recommended});$   
 $t'_{start} \leftarrow convert((head\ ordered'_L).timestamp);$   
 $grouped'_{launched} \leftarrow groupByTimeUnit(ordered'_L, t'_{start}, 0, timeUnit);$   
 $grouped'_{recommended} \leftarrow$   
 $\quad groupByTimeUnit(ordered'_R, t'_{start}, 0, timeUnit);$   
 $grouped'_{followed} \leftarrow [];$   
**foreach**  $G$  **in**  $grouped'_{launched}$  **do**  
 $\quad curGrouping \leftarrow [];$   
 $\quad$  **foreach**  $G_n$  **in**  $G$  **do**  
 $\quad\quad$  **if**  $G_n.context.statement \neq nil$  **then**  
 $\quad\quad\quad$  **do**  
 $\quad\quad\quad\quad curGrouping' \leftarrow curGrouping \cap G_n;$   
 $\quad\quad\quad\quad$  **recur**  $curGrouping'$   
 $\quad\quad$  **else**  
 $\quad\quad\quad$  **recur**  $curGrouping'$   
 $\quad\quad$  **end**  
 $\quad$  **end**  
 $\quad grouped'_{followed} \leftarrow grouped'_{followed} \cap curGrouping';$   
 $\quad$  **recur**  $grouped'_{followed}$   
**end**  
 $C_{launched} \leftarrow \text{map count}() grouped'_{launched};$   
 $C_{recommended} \leftarrow \text{map count}() grouped'_{recommended};$   
 $C_{followed} \leftarrow \text{map count}() grouped'_{followed};$   
 $combined \leftarrow [];$   
**for**  $i \leftarrow 0$  **to**  $count(C_{launched})$  **by** 1 **do**  
 $\quad tr_{starti} \leftarrow (first(nth(grouped'_{launched}, i))).timestamp;$   
 $\quad tr_{endi} \leftarrow (last(nth(grouped'_{launched}, i))).timestamp;$   
 $\quad N_{Li} \leftarrow nth(C_{launched}, i);$   
 $\quad N_{Ri} \leftarrow nth(C_{recommended}, i);$   
 $\quad N_{Fi} \leftarrow nth(C_{followed}, i);$   
 $\quad P_{Fi} \leftarrow N_{Fi} \div N_{Ri};$   
 $\quad P_{dueto_i} \leftarrow N_{Fi} \div N_{Li};$   
 $\quad subVec_i \leftarrow [tr_{starti}, tr_{endi}, N_{Li}, N_{Ri}, N_{Fi}, P_{Fi}, P_{dueto_i}];$   
 $\quad combined' \leftarrow combined \cap subVec_i$   
**end**  
**return**  $combined'$

---

- **map count() grouped'...** means apply the function **count()** to every sequence within the sequence *grouped'...* and put all results into a single array.

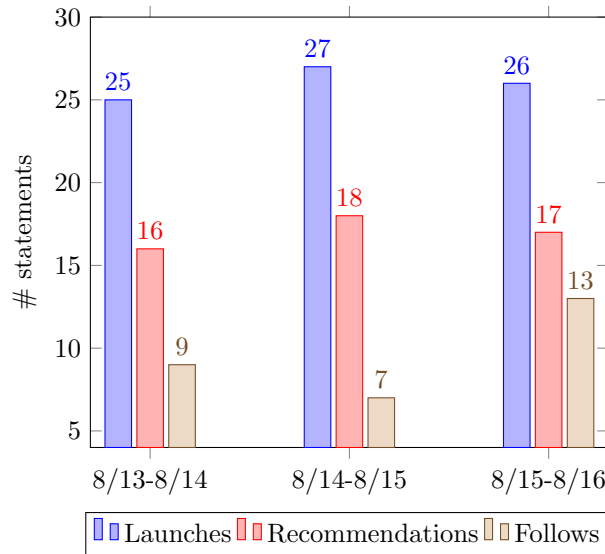
## 15.8 JSON Schema

```
{ "type": "array",  
  "items": { "type": "array",  
    "items": [ { "type": "string" }, { "type": "string" },  
               { "type": "number" }, { "type": "number" },  
               { "type": "number" }, { "type": "number" },  
               { "type": "number" } ] ] }
```

## 15.9 Visualization Description

The **Followed Recommendations** visualization can be a bar chart where the domain is time ranges and the range is a number representing the total count of statements recorded. For each time range, there will be three groups: 1) the number of launched statements 2) the number of recommended statements 3) the number of launches which are due to recommendations. Above each grouping or on hover, summary statistics can be displayed which describe the percentage of launches due to recommendations and the percentage of recommendations which were followed.

## 15.10 Visualization prototype



- The percentages described in section 5.9 are not displayed here.

## 15.11 Prototype Improvement Suggestions

Additional features may be implemented on top of this base specification but they would require adding additional values to each subarray returned by the

algorithm. These additional values can be retrieved via (1) performing metadata lookup within or independently of the algorithm (2) by utilizing additional xAPI statement parameters and/or (3) by performing additional computations. The following examples assume the metadata is contained within each statement available to the algorithm.

- populate a tooltip with the most popular launched, recommended and followed activity.
- populate a tooltip with the number of actors associated with the launches and follows.
- populate a tooltip with the actor who most often and the actor who least often follows recommendations.

## Appendix A: Visualization Exemplars

Appendix A includes a typology of data visualizations which may be supported within DAVE workbooks. These visualizations can either be one to one or one to many in regards to the algorithms defined within this document. Future iterations of this document will increasingly include these typologies within the domain-question template exemplars.

## Line Charts

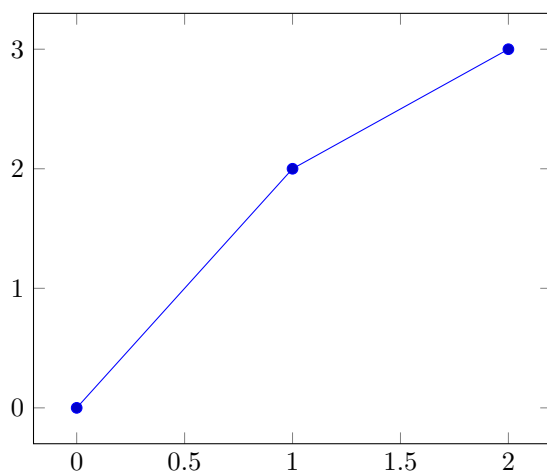


Figure 1: Line Chart

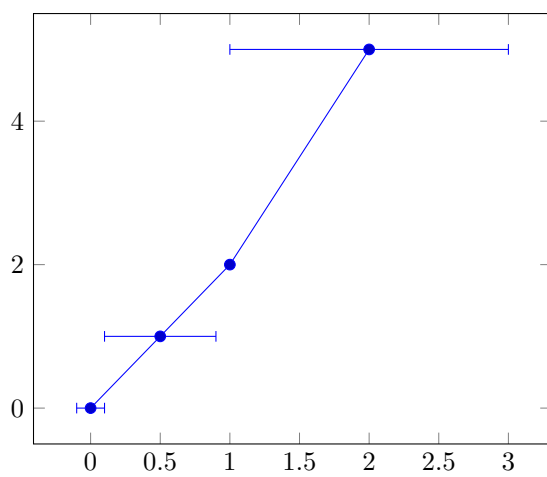


Figure 2: Line Chart with Error



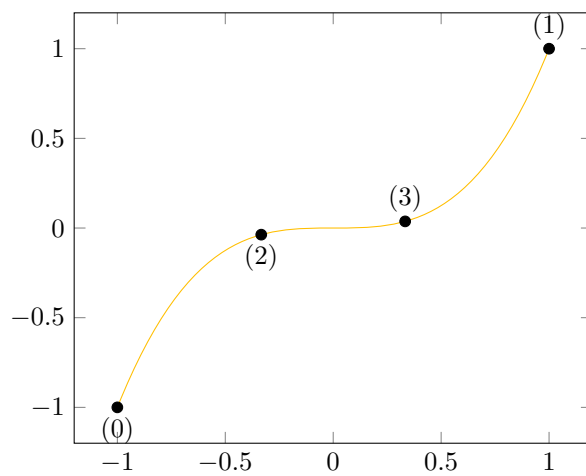


Figure 3: Spline Chart

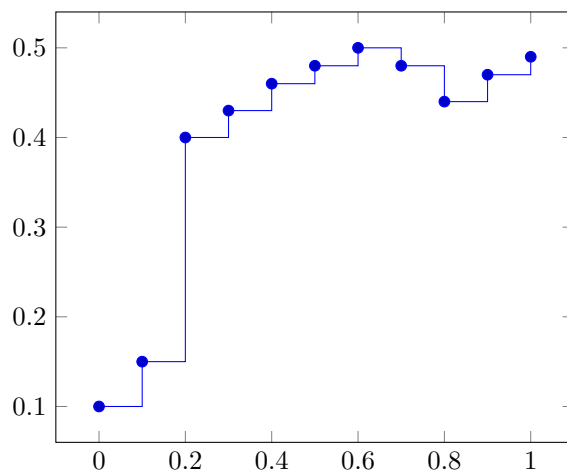


Figure 4: Quiver Chart

## Grouping Charts

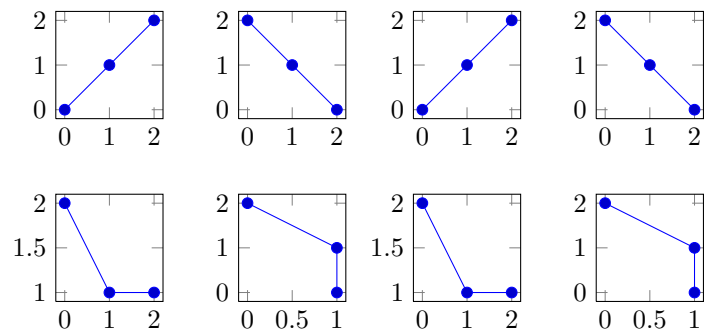


Figure 5: Grouped Line Charts

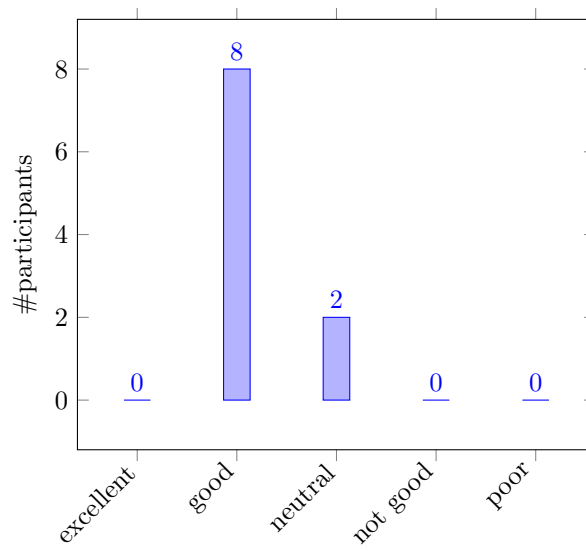


Figure 6: Histogram

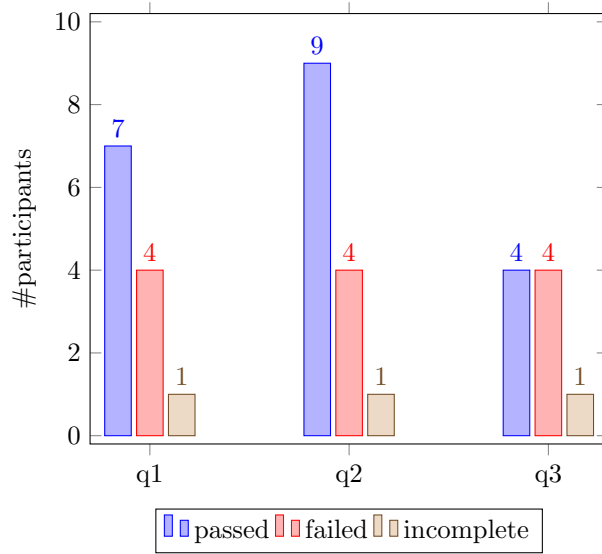


Figure 7: Bar Chart

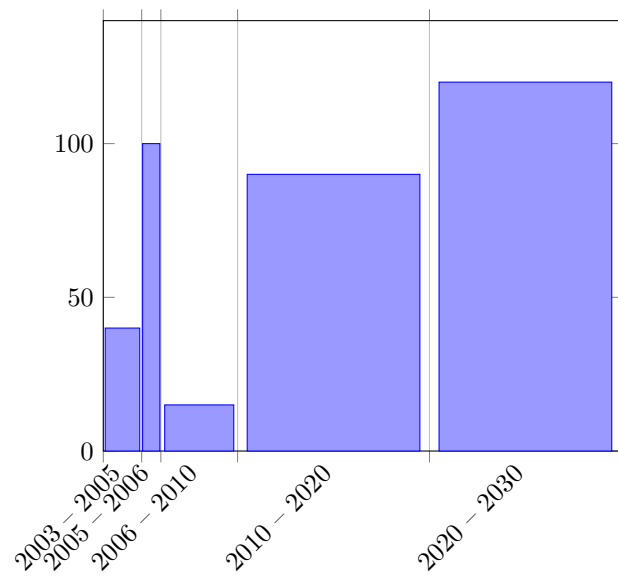


Figure 8: Bar Chart Grouped by Time Range

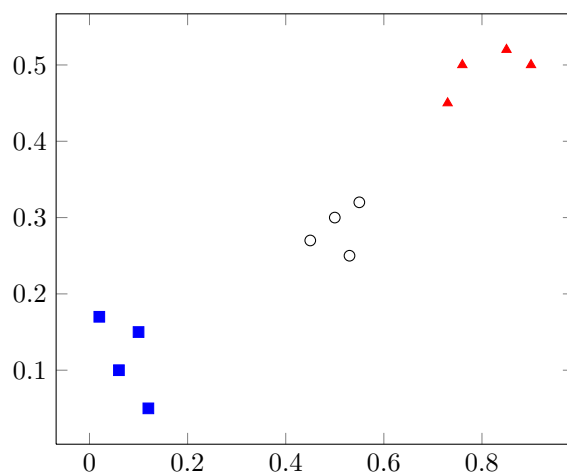


Figure 9: Scatter Plot

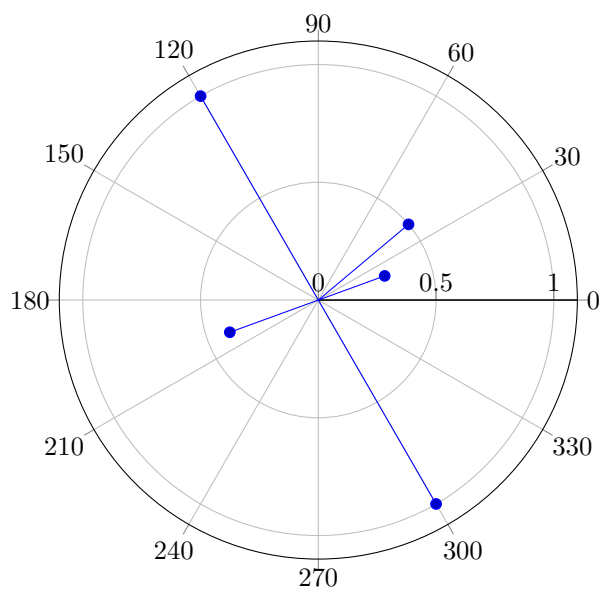


Figure 10: Polar Chart

## Specialized Charts

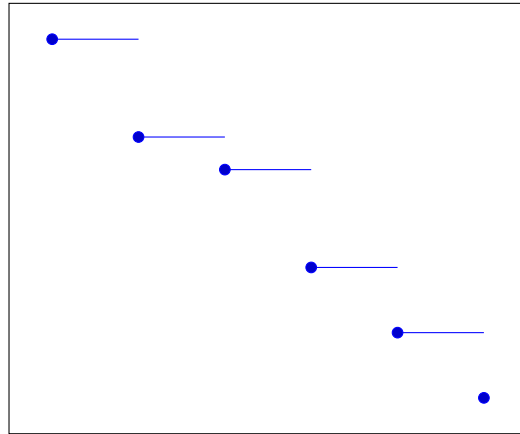


Figure 11: Gantt Chart

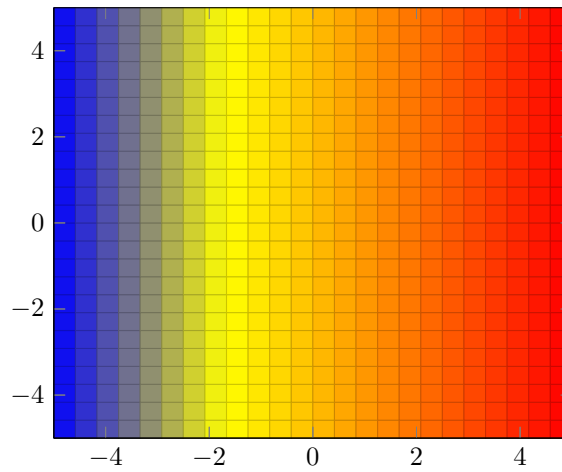


Figure 12: Heat Map

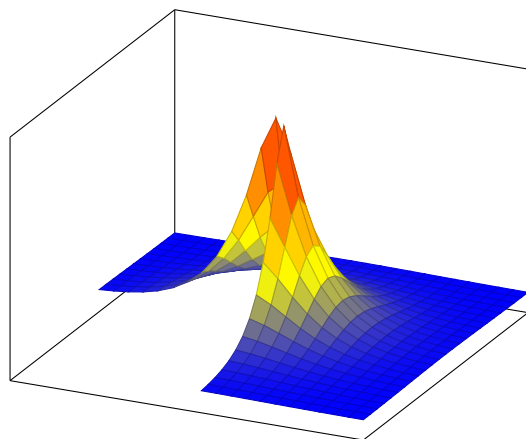


Figure 13: 3D Plot

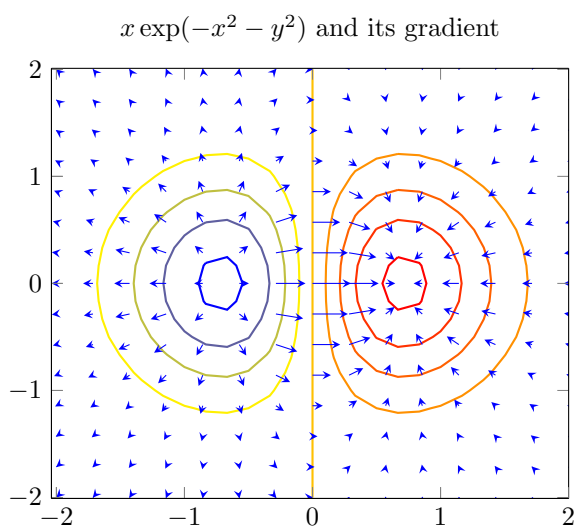


Figure 14: Gradient Plot