Learning restricted regular expressions with interleaving from XML data

1 Introduction for Inference Methods

IntelliJ IDEA An XSD (XML Schema Definition) Schema is required for running structure validation checks on a Web content file. IntelliJ IDEA can scan any XML file for the existing elements and attributes and generate a Schema for it.

Liquid Studio This is an XML Schema Generator tool, which uses a Wizard to create a compliant XML Schema by inferring its structure from sample XML documents. Configuration options allow a strict or lax approach to validation depending on the quality of your sample XML documents.

Altova XMLSpy The DTD/Schema or Generate DTD/Schema command generates a new DTD or W3C XML Schema from an XML document (or from a set of XML documents contained in a folder). This command is useful when you want to generate a DTD or XML Schema from XML documents.

Trang Trang, multi-format schema converter based on RELAX NG. It supports the following languages: RELAX NG (both XML and compact syntax), XML 1.0 DTDs, W3C XML Schema. It is implemented in Java and written by James Clark. Trang is an application for converting different XML schema languages to and from RELAX NG.

Trang can also infer a schema from one or more example XML documents which aims to produce human-understandable schemas. It tries to preserve all aspects of the input schema that may be significant to a human reader, including the definitions, the way the schema is divided into files, annotations and comments.

InstanceToSchema InstanceToSchema is a RELAX NG schema generator from XML instances. Implemented in Java. Written by Didier Demany. Open source (BSD license). InstanceToSchema is developed inside the xmloperator project and shares its license but is packaged and can be used independently from the XML editor.

A typical use case consists to obtain a description of the structure of one or several (combined) XML files. It must be noted that the tool make use of only a little part of the RELAX NG language.

Soa2Sore It is an algorithm for inferring subclass *SORE*.

Soa2Chare It is an algorithm for inferring subclass *Simplified CHARE*.

$Learn_{DME}^{+}$ It is an algorithm for inferring subclass *DME*.

conMiner It is an algorithm for inferring subclass *SIRE*.

GenEchare It is an algorithm for inferring subclass *eSimplified CHARE*.