

Table1 The original class XMLDTDValidator in Xerces

Original Class XMLDTDValidator	
Attributes	Methods
TOP_LEVEL_SCOPE	reset(XMLComponentManager)
NAMESPACES	getRecognizedFeatures()
VALIDATION	setFeature(String, boolean)
DYNAMIC_VALIDATION	getRecognizedProperties()
WARN_ON_DUPLICATE_ATTDEF	setProperty(String, Object)
SYMBOL_TABLE	setDocumentHandler(XMLDocumentHandler)
ERROR_REPORTER	setDTDHandler(XMLDTDHandler)
GRAMMAR_POOL	setDTDContentModelHandler(XMLDTDContentModelHandle
DATATYPE_VALIDATOR_FACTORY	r)
VALIDATION_MANAGER	startDocument(XMLLocator, String, Augmentations)
RECOGNIZED_FEATURES	xmlDecl(String, String, String, Augmentations)
RECOGNIZED_PROPERTIES	doctypeDecl(String, String, String, Augmentations)
DEBUG_ATTRIBUTES	startPrefixMapping(String, String, Augmentations)
DEBUG_ELEMENT_CHILDREN	startElement(QName, XMLAttributes, Augmentations)
ValidationManager	emptyElement(QName, XMLAttributes, Augmentations)
fValidationManager	characters(XMLString, Augmentations)
ValidationState fValidationState	ignorableWhitespace(XMLString, Augmentations)
fNamespaces	endElement(QName, Augmentations)
fValidation	endPrefixMapping(String, Augmentations)
fDTDValidation	startCDATA(Augmentations)
fDynamicValidation	endCDATA(Augmentations)
fWarnDuplicateAttdef	endDocument(Augmentations)
SymbolTable fSymbolTable	startExternalSubset(Augmentations)
XMLErrorHandler fErrorHandler	endExternalSubset(Augmentations)
DTDGrammarBucket	startGeneralEntity(String, XMLResourceIdentifier, String,
fGrammarBucket	Augmentations)
DTDDVFactory	checkStandaloneEntityRef(String)
fDatatypeValidatorFactory	comment(XMLString, Augmentations)
XMLDocumentHandler	processingInstruction(String, XMLString, Augmentations)
fDocumentHandler	endGeneralEntity(String, Augmentations)
XMLDTDHandler fDTDHandler	startDTD(XMLLocator, Augmentations)
XMLDTDCContentModelHandler	ignoredCharacters(XMLString, Augmentations)
fDTDCContentModelHandler	textDecl(String, String, Augmentations)
fDTDDGrammar	startParameterEntity(String, XMLResourceIdentifier, String,
fPerformValidation	Augmentations)
fSkipValidation	endParameterEntity(String, Augmentations)
fInDTDIgnore	elementDecl(String, String, Augmentations)
fCurrentElement	startAttlist(String, Augmentations)
fCurrentElementIndex	attributeDecl(String, String, String, String[], String, XMLString,
fCurrentContentSpecType	XMLString, Augmentations)
	endAttlist(Augmentations)

fRootElement	internalEntityDecl(String, XMLString, XMLString, Augmentations)
fSeenDoctypeDecl	
fInCDATASection	externalEntityDecl(String, XMLResourceIdentifier, Augmentations)
fElementIndexStack	
fContentSpecTypeStack	unparsedEntityDecl(String, XMLResourceIdentifier, String, Augmentations)
fElementQNamePartsStack	
fElementChildren	notationDecl(String, XMLResourceIdentifier, Augmentations)
fElementChildrenLength	startConditional(short, Augmentations)
fElementChildrenOffsetStack	endConditional(Augmentations)
fElementDepth	endDTD(Augmentations)
fSeenRootElement	startContentModel(String, Augmentations)
fInElementContent	any(Augmentations)
fMixed	empty(Augmentations)
fTempElementDecl	startGroup(Augmentations)
fTempAttDecl	pcdata(Augmentations)
fEntityDecl	element(String, Augmentations)
fTempQName	separator(short, Augmentations)
fNDataDeclNotations	occurrence(short, Augmentations)
fDTDElementDeclName	endGroup(Augmentations)
fMixedElementTypes	endContentModel(Augmentations)
fDTDElementDecls	addDTDDefaultAttrsAndValidate(int, XMLAttributes)
fBuffer	getExternalEntityRefInAttrValue(String)
fEMPTYSymbol	validateDTDattribute(QName, String, XMLAttributeDecl)
fANYSymbol	invalidStandaloneAttDef(QName, QName)
fCDATASymbol	normalizeAttrValue(XMLAttributes, int)
fIDSymbol	normalizeDefaultAttrValue(XMLString)
fIDREFSymbol	rootElementSpecified(QName)
fIDREFSSymbol	checkContent(int, QName[], int, int)
fENTITYSymbol	getContentSpecType(int)
fENTITIESSymbol	charDataInContent()
fNMTOKENSymbol	getAttributeTypeName(XMLAttributeDecl)
fNMTOKENSSymbol	init()
fNOTATIONSymbol	ensureStackCapacity(int)
fENUMERATIONSymbol	handleStartElement(QName, XMLAttributes)
fIMPLIEDSymbol	handleEndElement(QName, Augmentations, boolean)
fREQUIREDSymbol	
fFIXEDSymbol	
fValid	
fValidRef	
fValidRefs	
fValENTITY	
fValENTITIES	
fValNMTOKEN	
fValNMTOKENS	

fValNOTATION fTableOfIDAttributeNames fTableOfNOTATIONAttributeNames fNotationEnumVals	
---	--

Table2 The new classes extracted by developers

Class Extracted by the Developers - 1	
Attributes	Methods
TOP_LEVEL_SCOPE VALIDATION WARN_ON_DUPLICATE_ATTDEF SYMBOL_TABLE ERROR_REPORTER GRAMMAR_POOL RECOGNIZED_FEATURES RECOGNIZED_PROPERTIES fValidation fDTDValidation fWarnDuplicateAttdef SymbolTable fSymbolTable XMLErrorReporter fErrorReporter DTDGrammarBucket fGrammarBucket XMLDTDHandler fDTDHandler XMLDTDCContentModelHandler fDTDCContentModelHandler DTDGrammar fDTDGrammar fPerformValidation fInDTDIgnore fMixed XMLEntityDecl fEntityDecl fNDataDeclNotations fDTDElementDeclName fMixedElementTypes fDTDElementDecls fEMPTYSymbol fANYSymbol fCDATASymbol fIDSymbol fIDREFSymbol fIDREFSSymbol fENTITYSymbol fENTITIESSymbol fNMTOKENSymbol fNMTOKENSSymbol	setDTDHandler(XMLDTDHandler) setDTDCContentModelHandler(XMLDTDCContentModelHandle r) startExternalSubset(Augmentations) endExternalSubset(Augmentations) checkStandaloneEntityRef(String) startDTD(XMLLocator, Augmentations) ignoredCharacters(XMLString, Augmentations) startParameterEntity(String, XMLResourceIdentifier, String, Augmentations) endParameterEntity(String, Augmentations) elementDecl(String, String, Augmentations) startAttlist(String, Augmentations) attributeDecl(String, String, String, String[], String, XMLString, XMLString, Augmentations) endAttlist(Augmentations) internalEntityDecl(String, XMLString, XMLString, Augmentation) externalEntityDecl(String, XMLResourceIdentifier, Augmentations) unparsedEntityDecl(String, XMLResourceIdentifier, String, Augmentations) notationDecl(String, XMLResourceIdentifier, Augmentations) startConditional(short, Augmentations) endConditional(Augmentations) endDTD(Augmentations) startContentModel(String, Augmentations) any(Augmentations) empty(Augmentations) startGroup(Augmentations) pcdata(Augmentations) element(String, Augmentations) separator(short, Augmentations) occurrence(short, Augmentations) endGroup(Augmentations) endContentModel(Augmentations) normalizeDefaultAttrValue(XMLString)

fNOTATIONSymbol fENUMERATIONSymbol fIMPLIEDSymbol fREQUIREDSymbol fFIXEDSymbol fTableOfIDAttributeNames fTableOfNOTATIONAttributeNames fNotationEnumVals	
Class Extracted by the Developers - 2	
Attributes	Methods
fCDATASymbol fIDSymbol fIDREFSymbol fIDREFSSymbol fENTITYSymbol fENTITIESSymbol fNMTOKENSymbol fNMTOKENSSymbol fNOTATIONSymbol fENUMERATIONSymbol fIMPLIEDSymbol fREQUIREDSymbol fFIXEDSymbol TOP_LEVEL_SCOPE NAMESPACES VALIDATION DYNAMIC_VALIDATION WARN_ON_DUPLICATE_ATTDEF SYMBOL_TABLE ERROR_REPORTER GRAMMAR_POOL DATATYPE_VALIDATOR_FACTOR Y VALIDATION_MANAGER RECOGNIZED_FEATURES RECOGNIZED_PROPERTIES DEBUG_ATTRIBUTES DEBUG_ELEMENT_CHILDREN fValidationManager fValidationState fNamespaces fValidation fDTDValidation fDynamicValidation	reset(XMLComponentManager) getRecognizedFeatures() setFeature(String, boolean) getRecognizedProperties() setProperty(String, Object) setDocumentHandler(XMLDocumentHandler) startDocument(XMLLocator, String, Augmentations) xmlDecl(String, String, String, Augmentations) doctypeDecl(String, String, String, Augmentations) startPrefixMapping(String, String, Augmentations) startElement(QName, XMLAttributes, Augmentations) emptyElement(QName, XMLAttributes, Augmentations) characters(XMLString, Augmentations) ignorableWhitespace(XMLString, Augmentations) endElement(QName, Augmentations) endPrefixMapping(String, Augmentations) startCDATA(Augmentations) endCDATA(Augmentations) endDocument(Augmentations) comment(XMLString, Augmentations) processingInstruction(String, XMLString, Augmentations) startGeneralEntity(String, XMLResourceIdentifier, String, Augmentations) endGeneralEntity(String, Augmentations) textDecl(String, String, Augmentations) addDTDDefaultAttrsAndValidate(int, XMLAttributes) getExternalEntityRefInAttrValue(String) validateDTDattribute(QName, String, XMLAttributeDecl) invalidStandaloneAttDef(QName, QName) normalizeAttrValue(XMLAttributes, int) rootElementSpecified(QName) checkContent(int, QName[], int, int) getContentSpecType(int) charDataInContent()

fWarnDuplicateAttdef	getAttributeTypeName(XMLAttributeDecl)
fSymbolTable	init()
fErrorReporter	ensureStackCapacity(int)
fGrammarBucket	handleStartElement(QName, XMLAttributes)
fDatatypeValidatorFactory	handleEndElement(QName, Augmentations, boolean)
fDocumentHandler	normalizeDefaultAttrValue(XMLString)
fDTDGrammar	startExternalSubset(Augmentations)
fPerformValidation	endExternalSubset(Augmentations)
fSkipValidation	startAttlist(String, Augmentations)
fCurrentElement	startParameterEntity(String, XMLResourceIdentifier, String, Augmentations)
fCurrentElementIndex	endParameterEntity(String, Augmentations)
fCurrentContentSpecType	
QName fRootElement	
fSeenDoctypeDecl	
fInCDATASection	
fElementIndexStack	
fContentSpecTypeStack	
fElementQNamePartsStack	
fElementChildren	
fElementChildrenLength	
fElementChildrenOffsetStack	
fElementDepth	
fSeenRootElement	
fInElementContent	
fTempElementDecl	
fTempAttDecl	
fEntityDecl	
fTempQName	
fBuffer	
fValID	
fValIDRef	
fValIDRefs	
fValENTITY	
fValENTITIES	
fValNMTOKEN	
fValNMTOKENS	
fValNOTATION	

Table3 The new classes extracted by our approach

Class Extracted our approach - 1	
Attributes	Methods
fNDataDeclNotations	attributeDecl(String, String, String, String[],String, XMLString,
fDTDElementDecls	XMLString, Augmentations)
fDTDElementDeclName	endDTD(Augmentations)
fMixedElementTypes	setDTDHandler(XMLDTDHandler)

fMixed fDTDContentModelHandler fDTDHandler fDTDDGrammar fSkipValidation	startParameterEntity(String, XMLResourceIdentifier, String, Augmentations) startDTD(XMLLocator, Augmentations) elementDecl(String, String, Augmentations) notationDecl(String, XMLResourceIdentifier, Augmentations) unparsedEntityDecl(String, XMLResourceIdentifier, String, Augmentations) externalEntityDecl(String, XMLResourceIdentifier, Augmentations) internalEntityDecl(String, XMLString, XMLString, Augmentations) endConditional(Augmentations) startConditional(short, Augmentations) endParameterEntity(java.lang.String, org.apache.xerces.xni.Augmentations) endExternalSubset(org.apache.xerces.xni.Augmentations) endAttlist(org.apache.xerces.xni.Augmentations) comment(org.apache.xerces.xni.XMLString, org.apache.xerces.xni.Augmentations) startExternalSubset(XMLResourceIdentifier, Augmentations) startAttlist(String, Augmentations) ignoredCharacters(XMLString, Augmentations) element(String, Augmentations) startContentModel(String, Augmentations) setDTDContentModelHandler(XMLDTDDContentModelHandler) XMLDTDDProcessor.any(Augmentations) separator(short, Augmentations) occurrence(short, Augmentations) pcdata(Augmentations) startGroup(Augmentations) empty(Augmentations) endGroup(Augmentations) endContentModel(Augmentations)
Class Extracted our approach - 2	
Attributes	Methods
fEMPTYSymbol fANYSymbol fCDATASymbol fIDSymbol fIDREFSymbol fIDREFSSymbol fENTITYSymbol fENTITIESSymbol	endElement(QName, Augmentations) getContentSpecType(int) doctypeDecl(String, String, String, Augmentations) processingInstruction(String, XMLString, Augmentations) startGeneralEntity(String, XMLResourceIdentifier, String, Augmentations) comment(XMLString, Augmentations) characters(XMLString, Augmentations)

fNMTOKENSymbol	normalizeAttrValue(XMLAttributes, int)
fNMTOKENSSymbol	getAttributeTypeName(XMLAttributeDecl)
fNOTATIONSymbol	getExternalEntityRefInAttrValue(String)
fENUMERATIONSymbol	init()
fIMPLIEDSymbol	reset(parser.XMLComponentManager)
fREQUIREDSymbol	handleEndElement(QName, Augmentations, boolean)
fFIXEDSymbol	handleStartElement(QName, XMLAttributes, Augmentations)
fValidation	checkContent(int, QName[], int, int)
fDynamicValidation	addDTDDefaultAttrsAndValidate(QName, int, XMLAttributes)
fDatatypeValidatorFactory	ensureStackCapacity(int)
fDTDValidation	charDataInContent()
fWarnDuplicateAttdef	rootElementSpecified(QName)
fSymbolTable	validateDTDattribute(QName, String, XMLAttributeDecl)
fNamespaces	getRecognizedFeatures()
fBuffer	getRecognizedProperties()
fTempQName	invalidStandaloneAttDef(QName, QName)
fTempAttDecl	setFeature(String, boolean)
fEntityDecl	setProperty(String, Object)
fGrammarBucket	startDocument(XMLLocator, String, NamespaceContext, Augmentations)
fRootElement	startCDATA(Augmentations)
fInCDATASection	endCDATA(Augmentations)
fTempElementDecl	setDocumentHandler(XMLDocumentHandler)
fDTDDGrammar	startElement(QName, XMLAttributes, Augmentations)
fPerformValidation	emptyElement(QName, XMLAttributes, Augmentations)
fErrorReporter	endDocument(Augmentations)
fCurrentElement	endGeneralEntity(String, Augmentations)
fSeenDoctypeDecl	ignorableWhitespace(XMLString, Augmentations)
fCurrentElementIndex	textDecl(String, String, Augmentations)
fElementDepth	xmlDecl(String, String, String, Augmentations)
fContentSpecTypeStack	endPrefixMapping(String, Augmentations)
fElementIndexStack	startPrefixMapping(String, String, Augmentations)
fElementQNamePartsStack	
fElementChildrenOffsetStack	
fCurrentContentSpecType	
fElementChildren	
fElementChildrenLength	
fInElementContent	
fSeenRootElement	
fValidationManager	
fValidationState	
DATATYPE_VALIDATOR_FACTOR	
Y	
DEBUG_ATTRIBUTES	
DEBUG_ELEMENT_CHILDREN	

DYNAMIC_VALIDATION ERROR_REPORTER GRAMMAR_POOL NAMESPACES SYMBOL_TABLE TOP_LEVEL_SCOPE VALIDATION VALIDATION_MANAGER WARN_ON_DUPLICATE_ATTDEF fValENTITIES fValENTITY fValid fValidRef fValidRefs fValNMTOKEN fValNMTOKENS fValNOTATION RECOGNIZED_FEATURES RECOGNIZED_PROPERTIES fDocumentHandler	
---	--