Informatica

Informatica® Data Quality
10.5

Rule Specification Guide

Informatica Data Quality Rule Specification Guide
10.5
March 2021

© Copyright Informatica LLC 1998, 2021

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at https://www.informatica.com/trademarks.html. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2021-04-01

Table of Contents

Preface 6

Informatica Resources. 6

Informatica Network. 6

Informatica Knowledge Base. 6

Informatica Documentation. 6

Informatica Product Availability Matrices. 7

Informatica Velocity. 7

Informatica Marketplace. 7

Informatica Global Customer Support. 7

Chapter 1: Introduction to Rule Specifications. 8

Introduction to Rule Specifications. 8

Rule Specification Components. 9

Rule Sets. 10

Inputs. 10

Rule Statements. 10

Rule Specifications and Mapplets. 11

Rules and Guidelines for Rule Specifications and Mapplets. 11

Rule Specifications and Version Control. 12

Working with Rule Specifications in a Versioned Model Repository. 13

Chapter 2: Rule Specification Configuration. 14

Rule Specification Configuration Overview. 14

Steps to Configure a Rule Specification. 14

Verifying the Business Rule Requirements. 15

Verifying the Business Data Properties. 15

Designing the Rule Specification. 15

Rule Specification Properties. 16

Rule Specifications and Business Glossary Terms. 17

Configuring the Rule Specification. 17

Creating an Input. 18

Creating a Rule Specification from a Business Term. 18

Generating a Mapplet from a Rule Specification. 19

Rules and Guidelines for Generating Mapplets. 19

Opening a Rule Specification. 20

Chapter 3: Rule Set Configuration. 21

Rule Set Configuration Overview. 21

Rules and Guidelines for Rule Set Configuration. 21

Rule Set Configuration Examples. 22

Rule Sets that Read a Single Data Input. . . . 22
Rule Sets with Dependent Conditions. . . . 23
Rule Sets with Independent Conditions. . . . 23
Parent Rule Sets and Child Rule Sets. . . . 23
Rule Set Properties. . . . 24
Adding a Rule Set to a Rule Specification. . . . 25
Cut, Copy, and Paste Operations on Rule Sets. . . . 25
Copying and Moving a Rule Set in a Rule Specification. . . . 25
Copying and Moving a Rule Set to a Different Rule Specification. . . . 26

Chapter 4: Rule Statement Configuration. . . . 27

Rule Statement Configuration Overview. . . . 27
Conditions. . . . 28
Function Configuration in Conditions. . . . 29
Operators. . . . 30
Actions. . . . 31
Function Configuration in Actions. . . . 33
Function Expression Descriptions. . . . 34
Date Elements in Functions . . . 36
Reference Tables. . . . 36
Rule Statement Menu Options. . . . 37
Configuring a Rule Statement. . . . 38
Configuring a Rule Statement that Reads the Result of Another Rule Statement. . . . 39
Configuring a Rule Statement with Multiple Conditions. . . . 39
Cut, Copy, and Paste Operations on Rule Statements. . . . 40
Copying and Moving a Rule Statement in a Rule Specification. . . . 40
Copying and Moving a Rule Statement to a Different Rule Specification. . . . 40

Chapter 5: Common Types of Rule Statements. . . . 42

Common Types of Rule Statements Overview. . . . 42
Verifying the Accuracy of Business Data . . . 44
Identifying Obsolete Values in Business Data. . . . 44
Standardizing Business Data Values. . . . 45
Improving the Usability of Business Data . . . 46
Discovering Information About Business Users. . . . 47
Using Data Values to Verify a Business Policy. . . . 48
Updating Data Values to Meet Business Standards. . . . 49
Verifying the Application of a Business Rule. . . . 49
Validating Address Records. . . . 50
Using a List of Values to Find Records in a Data Set. . . . 52
Identifying Data Values that Contain Key Words or Character Strings. . . . 53

Chapter 6: Test and Validation Operations. 55

Test and Validation Operations Overview. 55

Rules and Guidelines for Validation Operations. 55

Validating a Rule Specification. 56

Testing a Rule Specification. 56

Testing a Rule Set. 57

Index. 58

Preface

Read the *Rule Specification Guide* to learn how to configure rule specifications in Informatica Analyst. A rule specification is an object that represents the data requirements of a business rule in logical form. You can add a rule specification to a mapping in Informatica Developer to verify that the source data in the mapping conforms to the criteria that the business rule specifies. You can also generate mapplets from a rule specification and use the mapplets in mappings in Informatica Developer.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit https://network.informatica.com.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit https://search.informatica.com. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit https://docs.informatica.com.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at https://network.informatica.com/community/informatica-network/product-availability-matrices.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at http://velocity.informatica.com. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at https://marketplace.informatica.com.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:
https://www.informatica.com/services-and-training/customer-success-services/contact-us.html.

To find online support resources on the Informatica Network, visit https://network.informatica.com and select the eSupport option.

CHAPTER 1

Introduction to Rule Specifications

This chapter includes the following topics:

- Introduction to Rule Specifications, 8
- Rule Specification Components, 9
- Rule Specifications and Mapplets, 11
- Rule Specifications and Version Control, 12

Introduction to Rule Specifications

A rule specification is an asset that represents the data requirements of a business rule in logical form. You configure a rule specification in the Design workspace of the Analyst tool.

Use a rule specification to define the following data operations:

- Define the types of data that a business data set contains.
- Define a set of conditions that the business data must satisfy.
- Define the actions to take when the data satisfies the conditions of the business rule.
- Define the actions to take when the data fails to satisfy the conditions of the business rule.

When you complete work on a rule specification, you can generate a metadata object called a mapplet from the rule specification. The mapplet contains the same business logic as the rule specification. An Informatica Developer user can connect the rule specification or the corresponding mapplet to a data source to verify that the data conforms to the business rule. A Developer tool user can edit a mapplet but cannot edit a rule specification.

A rule specification or mapplet applies the conditions and actions that you define to the values in each row in a data source. The rule specification or mapplet generates an output value for each row. The output specifies whether the row meets the requirements that you define for the business rule.

Related Topics:

- "Rule Specifications and Mapplets" on page 11

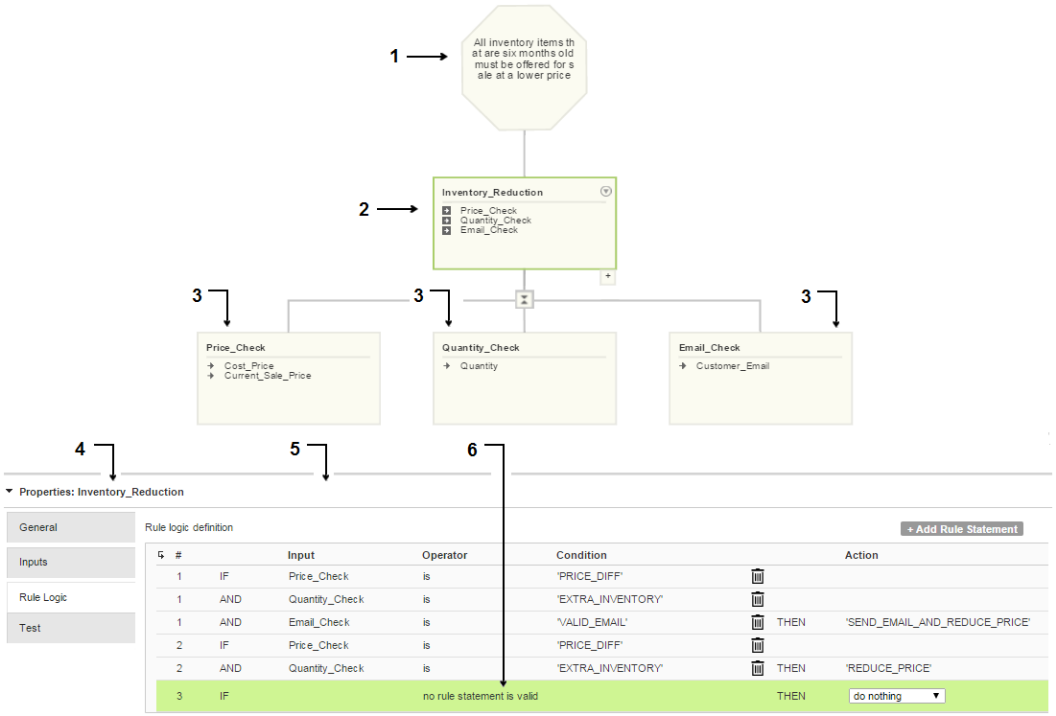
Rule Specification Components

When you create a rule specification, you configure a series of shapes in the Design workspace. Each shape describes an aspect of the business rule. The top-level shape stores the name and the general properties of the rule specification. The other shapes describe the business rule logic.

The shapes below the top-level shape are rule sets. A rule set contains rule statements that define the business rule requirements at a low level. Each rule statement reads a column of input data and verifies that the input data meets the conditions that you specify.

The rule set that connects to the top-level shape is the primary rule set. The primary rule set summarizes the business rule. The output from the primary rule set determines whether each row of input data meets the requirements of the business rule.

The following image shows a rule specification in the Design workspace:



The rule specification contains the following elements:

1. Top-level shape.
2. Primary rule set.
3. Child rule sets of the primary rule set.
4. Property options on the rule set that you select.
5. Rule statements in the rule set that you select.
6. Rule statement that the system defines.

Rule Sets

Rule sets define the logical flow of data through the rule specification. Data flows upward through the rule specification from the lowest rule set to the primary rule set.

You can add a rule set below any rule set in the rule specification. The rule sets have a parent-to-child relationship. When you add a rule set, the output of the rule set that you add becomes an input to the parent rule set.

Use rule sets to define the rule statements that analyze and update the input data. You can configure a rule set with a single rule statement, or you can add multiple rule statements to the rule set. Within a rule set, data flows from the first rule statement to the final rule statement.

You can copy or move a rule set to another location in a rule specification, and you can copy or move a rule set to another rule specification.

Inputs

An input describes a column of data that a rule statement can analyze. An input represents a column in a data set, or an input represents the output from another rule set in the rule specification.

Use the **Manage Global Inputs** dialog box to define an input.

When you define an input, specify the following properties:

- The data type of the data that the input represents. Create an input with a date/time, float, integer, or string data type.

  Note: You can specify an integer data type for numbers in the range -2147483648 through 2147483647. To read numbers that are outside the integer range, use the float data type.
- The maximum number of characters that a value in the column can contain.

You can optionally enter a text description of the input.

An input does not store information about the business data, such as the name of a column, table, or database. The Developer tool user adds the mapplet or the rule specification to a mapping and connects the mapping inputs to the business data. You advise the developer about the columns to connect. The developer can connect an input to any column that matches the properties that you define.

When you create a child rule set, the output from the rule set becomes an input to the parent rule set. You must use the input in a rule statement in the parent rule set. When you create an input in the **Manage Global Inputs** dialog box, you can optionally add the input to a rule statement. To view the inputs that a rule set uses, select the rule set in the workspace.

Rule Statements

A rule statement is an IF-THEN statement that analyzes a column of data and generates an output based on the result of the analysis. You add a rule statement to a rule set.

A rule statement uses conditions and actions to define the IF-THEN logic. A condition is a data operation that determines a single fact about a data value. You can add multiple conditions to a rule statement.

An action is a data operation that generates the output from the rule set. An action generates data when the input that you add to the rule statement satisfies the conditions that you define. The rule set uses the output from the first rule statement that generates output data.

Each rule set contains a system-defined rule statement that specifies the action to take if no other rule statement generates output data. The rule statement is the final rule statement in the rule set. You can edit the action in the system-defined rule statement. By default, the rule statement specifies that the rule set does not generate any output data if the other rule statements do not generate output data.

Rule Specifications and Mapplets

Developer tool users can use the rule specifications that you create. The users can also use any mapplet that you generate from a rule specification. Rule specifications are read-only in the Developer tool, whereas Developer tool users can edit mapplets.

When you generate mapplets, the Analyst tool generates a mapplet for each rule set in a rule specification. The mapplet that corresponds to the primary rule set contains the complete logic of the rule specification and has the same name as the rule specification.

A Developer tool user can add a rule specification or a mapplet to a mapping. A mapping is a metadata object that connects to a data source and writes to a target data object. Informatica stores mapplets and rule specifications in the Model repository. The Model repository is the storage database for mappings, rule specifications, mapplets, and other assets in the Informatica domain. Add a rule specification to a mapping when you want to apply the logic in the current rule specification to the data source. Add a mapplet that you generate to a mapping when you want to update the mapplet logic before you run the mapping.

A mapplet contains transformations. A transformation defines an operation to analyze or update data. Each rule statement in a rule specification defines at least one transformation in a mapplet.

A mapplet can contain other mapplets. If you define a complex rule specification, the mapplet that you generate contains multiple mapplets in a hierarchal structure. Each mapplet is a separate object in the Model repository. Because a mapplet can contain other mapplets, you can add a mapplet to a rule statement in a rule specification. The mapplet that you generate from the rule specification contains a reference to any mapplet that you added to a rule statement.

Add a mapplet to a rule statement in the following cases:

- The mapplet contains complex function logic. You use the mapplet so that you do not have to define the logic in the current rule specification.
- The mapplet contains function logic that you cannot configure in the Analyst tool. For example, the mapplet contains address validation logic.
- The mapplet contains function logic that the organization approves for all Analyst tool users.

After a mapping runs, the developer can send you the data output from the mapping. You can evaluate the data output to determine if the data source conforms to the business rule that the rule specification describes. Alternatively, the developer or another user can run a profile on the mapping output. A profile is an asset that describes the data patterns in a data set. You can evaluate the profile to determine the conformity of the business data to the business rule. You can also create and run a profile in the Analyst tool.

Rules and Guidelines for Rule Specifications and Mapplets

A key purpose of a rule specification is to define one or more mapplets that an Informatica developer can apply to the business data. You generate the mapplets from the rule specification. You can also add another mapplet to the rule specification that you create.

Consider the following rules and guidelines when you work with mapplets and rule specifications:

- A mapplet is a reusable object. A Developer tool user can add a mapplet to multiple mappings. You can add a mapplet to rule statements in multiple rule specifications.

  A rule specification is also a reusable object. A Developer tool user can add a rule specification to multiple mapplets and mappings. You cannot add a rule specification to a rule statement.
- When you generate mapplets from a rule specification, you create a mapplet for every rule set in the rule specification. Each mapplet appears as a unique object in the Model repository.
- The mapplet that represents the primary rule set takes the name of the rule specification.

A mapplet that represents a rule set below the primary rule set uses the following naming convention:

[rule specification name]_[rule set name]

Each rule set in the rule specification must have a unique name.

- A mapplet that represents a rule set includes logic for any rule set that creates an input below the current rule set. A mapplet that represents the primary rule set contains all of the rule specification logic.
- When you generate mapplets from a rule specification, the rule specification and the mapplet that represents the primary rule set are functionally identical. A Developer tool user can edit a mapplet that you generate from a rule specification. The rule specification object is a read-only object in the Developer tool and does not change unless you update the rule specification in the Analyst tool.
- You can add a rule specification to a profile in the Analyst tool.

  A Developer tool user can export a mapping that contains a rule specification to PowerCenter®. A Developer tool user can also deploy a rule specification as a web service.
- You can push down a mapping that contains a rule specification to a Hadoop cluster.

  You can push down a mapping that contains a mapplet that you generated from a rule specification to a Hadoop cluster.
- To view the list of mapplets that you generated in the Analyst tool, select the **Rules** option from the Assets view in the library workspace. A rule is a mapplet that you can use with other assets in the Analyst tool. The mapplets that you create appears as rules in the library.

  To a Developer tool user, rules and mapplets are the same type of object.
- A Developer tool session uses the rule specification objects that are present in the Model repository at the time that the Developer tool user connects to the repository. If you change a rule specification in the Analyst tool, the change is not available to the Developer tool session until the Developer tool user refreshes the repository connection.
- If you add, delete, or modify an input in a rule specification, you invalidate every link between the rule specification inputs and other objects in a mapplet or mapping.

  If you add, delete, or modify an output in a rule specification, you invalidate every link between the rule specification outputs and other objects in a mapplet or mapping.

  Modifications that break the links include changes to the name, precision, or data type of an input or output.
- If you update the business logic in a rule specification but you do not modify an input or output, you do not break any link to another object in a mapplet or mapping.
- The Analyst tool and the Developer tool use application services to read and write objects in the Model repository and to run mappings. The Analyst tool uses the Analyst Service to read and write rule specifications, reference tables, and mapplets. The Developer tool uses the Data Integration Service to run mappings.

  If you cannot perform an operation in a rule specification, ask an administrator to verify your permissions on the application services.

Rule Specifications and Version Control

If the Model repository that stores the rule specifications integrates with a version control application, you can apply version control to the rule specifications.

You can check in and check out rule specifications from a Model repository that supports version control. You can undo a checkout, retrieve an earlier version, and restore a rule specification to an earlier version.

When a rule specification is not under version control, the Model repository locks a rule specification while you edit it.

You can add a mapplet to a rule statement regardless of the versioned status of the mapplet. The rule statement reads the most recent version of the mapplet in the Model repository.

The rule specification does not store version information for a mapplet that you add to a rule statement. When you generate a mapplet from a rule specification, the mapplet that you generate includes the most recent version of any mapplet that you add.

Working with Rule Specifications in a Versioned Model Repository

To open the rule specification in read-only mode, click a rule specification name. To work on the rule specification, enter edit mode. If the rule specification is checked into the Model repository, check out the rule specification before you edit it.

1. On the Informatica toolbar, click **Open**.

   The asset library opens.

2. Select the Rule Specifications asset category, and select a rule specification name.

   The rule specification opens in read-only mode.

3. To edit the current version of the rule specification, click **Edit**.

   To edit the rule specification that a user checked in to a versioned Model repository, check out the rule specification.

   Note: If you right-click the rule specification name, you can use the menu options to open, edit, or check out the rule specification.

4. When you complete work on the rule specification, click **Save and Finish**.

   The Analyst tool saves your changes to the rule specification.

   If you checked out the rule specification from a versioned Model repository, check in the object. A versioned Model repository does not update the rule specification version until you check in the object.

CHAPTER 2

Rule Specification Configuration

This chapter includes the following topics:

- Rule Specification Configuration Overview, 14
- Steps to Configure a Rule Specification, 14
- Rule Specification Properties, 16
- Rule Specifications and Business Glossary Terms, 17
- Configuring the Rule Specification, 17
- Creating an Input, 18
- Creating a Rule Specification from a Business Term, 18
- Generating a Mapplet from a Rule Specification, 19
- Opening a Rule Specification, 20

Rule Specification Configuration Overview

When you configure a rule specification, you define a series of inputs, rule sets, and rule statements. Each element in the rule specification represents a data operation that interacts with the business data to validate a business rule.

Before you configure the rule specification, verify that you understand the properties of the business data and the requirements of the business rule.

Steps to Configure a Rule Specification

When you configure a rule specification, you translate the requirements of a business rule into one or more rule statements. The rule statements represent the logic that determines whether a data set conforms to the business rule. You use the rule specification to create one or more mapplets that a developer can apply to a data set.

Before you configure the rule specification in the Analyst tool, perform the following steps:

1. Verify the business rule requirements.
2. Verify the business data properties.
3. Determine the sequence of the rule statements.

Verifying the Business Rule Requirements

Before you configure a rule specification, discuss the business rule with the data owners in the organization. Verify that the business rule is valid and ready to apply to the business data.

1. Identify the business rule that the rule specification represents.
2. Identify the business data set that the business rule validates.
3. List the business rule requirements that apply to the data inputs.

   The business rule requirements indicate the types of rule statement that you create in the rule specification.
4. Identify the information types that the business rule applies to.

   The information types indicate the data types of the inputs that you create in the rule specification.

Verifying the Business Data Properties

Before you create a rule specification, identify the business data sets that the rule specification can apply to. Work with a developer to identify the data sets. When you generate a mapplet from the rule specification, the developer adds the mapplet to a mapping and connects the mapping to the data sets.

1. Identify one or more data sets that the developer can select as a data source in a mapping. For example, identify the database and the table that contain the data.
2. Verify the data types of the data columns that the mapplet analyzes. You specify the data types when you create the inputs in the rule specification.

   Note: You might not create an input for every column in the data set.
3. Discuss the steps to follow after the developer runs a mapping that contains the mapplet.

   The developer gives you the results of each mapping. You determine if you need to update the rule specification and generate the mapplet again. Alternatively, you determine if the business needs to update the business data so that it conforms to the business rule.

Designing the Rule Specification

At a high level, a business rule defines a single objective that the business data must satisfy. Add a rule statement that analyzes the objective to the primary rule set. If the business rule defines more than one data outcome, you might define more than one rule statement in the primary rule set.

Add rule sets and rule statements below the primary rule set to validate the data that the primary rule set analyzes.

1. Create the inputs that the rule statements analyze. The inputs are a key prerequisite for the rule statements. The inputs represent the columns in the business data set that the business rule applies to.
2. Identify the business rule conditions that the business data must satisfy. Plan a rule statement for each condition.
3. Determine the sequence in which the rule statements must analyze the data.

   Select the lowest dependencies in the business rule, and add a rule set for each dependency.
4. Add the rule sets that you need to the rule specification.

   Work from the lowest level in the rule specification upward to the primary rule set.
5. Add the rule statements that you identified to the rule sets.

Note: You can move a rule statement from one rule set to another rule set. You can move a rule set to another location in the rule specification, and you can move a rule set to another rule specification.

6. Test the rule specification.

   If the test data passes through the rule specification in the way that you expect, the rule specification is ready to use.

Rule Specification Properties

The rule specification properties define the behavior of the rule specification and identify the assets that the rule specification uses. To view the properties, click the top-level shape in the rule specification.

The rule specification displays the properties on the following views:

General

Lists the name, the description, and the Model repository location of the rule specification. You can update the name and the description.

Assets

Lists the assets that link to the rule specification. The rule specification can link to mapplets and to reference tables.

Click an asset name to open the asset in the workspace. Expand an asset to view the asset metadata.

The mapplets that you generate from the rule specification appear in the **Generated Assets** area. Any mapplet or reference table that you select in a rule statement appears in the **Related Assets** area.

Rule Properties

Lists the properties that the rule specification applies to the output from each rule set.

The rule specification applies the following properties to the output data:

- Maximum string length. Determines the maximum length for text data. Default is 100.
- Maximum number length. Determines the maximum length for numeric data. Default is 10.
- Number of decimal places. Determines the maximum number of decimal places to include in the output data. Default is 4.

  Note: A rule set can read a maximum of four decimal places from input data of the float data type.

Time Period

Specifies the time period in which the mapplets that you generate from the rule specification can run in a mapping. The properties apply to all mapplets that the rule specification generates, including any mapplet that derives from a rule set or a rule statement within the rule specification. The properties do not apply to a mapplet that you select in a condition or an action. By default, the rule specification does not specify a valid time period.

If you run a mapping that reads a mapplet outside the valid time period, the mapping fails to run. You can validate a rule specification and generate a mapplet rule from a rule specification outside the time period.

The properties also apply to the test operations that you perform on the rule specification. You cannot test a rule specification outside the time period.

Test

Displays the inputs that you add to the rule sets as a set of fields that you can populate with test data. Use the fields to verify that data flows through the rule specification in the ways that you expect. You can test the rule specification, and you can test any rule set in the rule specification.

To test a rule specification or a rule set, enter sample data to the fields in the test area. You can cut, copy, and paste data between test areas in the rule specification and between two rule specifications that are open concurrently. When you save the rule specification, you save the sample data that you entered.

To update the properties, open the rule specification and click **Edit**.

Rule Specifications and Business Glossary Terms

You can create a rule specification from a rule in a business term. A business term is a word or phrase that defines a business concept for the members of an organization. You can build a glossary of business terms in the Analyst tool.

You can add a rule to a business term. The rule specifies the data conditions that the business term must satisfy to be valid in the organization. For example, you might define a rule for the business term "Mortgage" in a financial business glossary. The rule defines a mortgage as a set of data values that the mortgage document must contain, such as the interest rate on the loan. You can create a rule specification from the rule and link the rule specification to the rule.

Use the **Assets** options in the business term to open the Design workspace for rule specifications. Create a rule specification, and define a rule statement for each data condition that the business term specifies.

Configuring the Rule Specification

After you analyze the business requirements and the business data, you can configure the rule specification.

1. Select Rule Specification from the **New** menu.
2. Enter a name for the rule specification.
3. Optionally, enter a description.

   Tip: Enter the business rule as the description.
4. Select a location for the rule specification.

   The Location field specifies a project in the Model repository.
5. Click **Continue**.

   The rule specification opens in the Design workspace.
6. Verify the general properties.
   - Select the top-level shape in the rule, and select **Rule Properties**.
   - Verify that the properties are correct for the data inputs that you create.
7. Optionally, set a valid time period for the mappings that you generate from the rule specification.
8. Configure a rule statement in the primary rule set.

The primary rule set defines the data output from the rule specification.

9. Optionally, update the primary rule set name.
   - Select the primary rule set, and select **General**.
   - Update the rule set name.
10. Add any rule set that the rule specification requires.

    You add a rule set below another rule set.
11. Add one or more inputs to each rule set.
    - Use the **Manage Global Inputs** dialog box to create the inputs.
12. Add any rule statement that the rule specification requires.

    You can add a rule statement to the primary rule set or to another rule set.
13. Save the rule specification.

After you configure the rule specification, test the rule specification with sample data.

Creating an Input

When you create an input, you can add the input to any rule set in the rule specification.

1. Open the rule specification.
2. Select the **Manage Global Inputs** option from the toolbar.

   The **Input Management** dialog box opens. The dialog box lists the current inputs in the rule specification.
3. Click **Add Input**.
4. Set the following properties on the input:
   - The input name. You can enter alphanumeric characters and underscore characters. Do not add spaces to the name.
   - The data type of the column in the business data set that the input represents. You can select a date/time, float, integer, or string data type.
   - The maximum character length of the data values in the input data column.
   - A text description of the input. You can enter a description of the data column. The description is optional.
5. To save the input and close the dialog box, click **OK**.

Note: The **Input Management** dialog box includes a Usage column. The Usage column indicates the number of times that each input appears in a rule statement in the rule specification.

Creating a Rule Specification from a Business Term

You can create a rule specification from a business term that specifies a rule.

1. Open the library.
2. Select **Business Terms** from the list of assets.

The library shows the business terms in the Model repository.

3. Select a business term.

   The business term opens in the Glossary workspace.

4. Click **Edit**.
5. Browse the business term properties, and find the rule from which to create the rule specification.
6. Click the option to link the rule to a rule specification.

   The Analyst tool opens the Design workspace for rule specifications.

   The Design workspace uses the following information from the rule in the business term:

   - The workspace displays the rule name as the default rule specification name.
   - The workspace displays the rule intent as the default rule specification description.

7. Save the rule specification. Optionally, continue to edit the rule specification.

Generating a Mapplet from a Rule Specification

When you finish work on a rule specification, you can generate one or more mapplets from the rule specification. The operation generates a mapplet for each rule set in the rule specification. The mapplet that the rule specification generates for the primary rule set represents the complete logic of the rule specification. A developer can use any mapplet that you generate in a mapping.

1. Open the rule specification.

   If the rule specification is read-only, click **Edit**.

   If the rule specification is under version control, check out the rule specification from the Model repository.

2. Validate and save the rule specification.

   If the rule specification contains a validation error, fix the error.

3. Select a location for the mapplets that you generate. You create the mapplets in a Model repository project.
4. Click **Generate rule**.

   Note: Mapplet generation discards any input that the rule specification logic does not use. If the rule specification contains an unused input, the Analyst tool displays a warning.

   The Analyst tool creates one or more mapplets in the Model repository.

Rules and Guidelines for Generating Mapplets

The **Generate rule** option creates one or more mapplets in the Model repository.

Consider the following rules and guidelines when you generate mapplets:

- If you update a rule specification after you generate mapplets, generate the mapplets again to update the mapplets in the Model repository.
- If you generate a mapplet in a Model repository under version control, the operation checks in the mapplet to the Model repository.

- If you generate a mapplet more than once in the same Model repository location, you replace the mapplet in the Model repository. If you replace a mapplet that links to other objects in a mapping, the operation deletes the port links on the mapplet in the mapping.

  If the mapplet is checked out of the Model repository, you cannot generate the mapplet to the same location. For example, the mapplet might be checked out in the Developer tool. Verify with the Developer tool user that you can check in the mapplet, and then try to generate the mapplet again. Or, generate the mapplet to another location in the Model repository.
- If you change the name of a rule set and you generate the mapplet, you create a mapplet with the current rule set name. The operation does not rename any earlier mapplet that you generated from the rule set.

Opening a Rule Specification

Open a rule specification from the Analyst tool library. You open a rule set in read-only mode. You can validate and test a rule specification in read-only mode. To update the rule specification or to generate a mapplet from the rule specification, open the rule specification in edit mode.

1. Open the library.
2. Select **Rule Specifications** from the list of assets.

   The library shows the rule specifications in the Model repository.
3. Select a rule specification name.

   The rule specification opens in the Design workspace.
4. To update the rule specification or to generate a mapplet, click **Edit**.

   If the rule specification is under version control, you must check out the rule specification before you enter edit mode.

When you finish work on the rule specification, click **Save and Finish**. If you checked out the rule specification from the Model repository, check in the rule specification.

CHAPTER 3

Rule Set Configuration

This chapter includes the following topics:

- Rule Set Configuration Overview, 21
- Rules and Guidelines for Rule Set Configuration, 21
- Rule Set Configuration Examples, 22
- Rule Set Properties, 24
- Adding a Rule Set to a Rule Specification, 25
- Cut, Copy, and Paste Operations on Rule Sets, 25

Rule Set Configuration Overview

You can configure a single rule set in a rule specification, or you can configure multiple rule sets. The number of rule sets that you configure depends on the facts that you need to verify about the business data.

A fact is a piece of information that you derive from a data value when you compare it to other data values. For example, a date value might indicate a birthday, or it might indicate the date that you sent an invoice to a customer. You can use the date to determine if a person is an adult, or if a customer account is overdue. You can use the facts that you determine about the person or about the customer to make business decisions. A rule specification can determine a single fact, or it can determine multiple facts that operate in a dependent relationship to each other.

When one fact about the business data depends on multiple other facts, create rule sets in a parent and child relationship. Use the child rule sets to generate inputs that the parent rule set can verify. If a fact relates to a single item of data, you can configure rule statements to analyze the data in a single rule set. If a fact depends on multiple independent facts about the business data, create child rule sets in parallel to determine each fact.

Rules and Guidelines for Rule Set Configuration

When you configure rule sets in the workspace, you define the flow of data through the rule specification. The position of the rule sets and the order of the rule statements in the rule sets determine the data flow.

Consider the following rules and guidelines when you configure rule sets:

- Data flows upwards through a rule specification from the lowest rule set to the primary rule set. When the rule specification includes parallel rule sets, the rule specification runs the parallel rule sets concurrently.

- A rule set generates a single output for a row of input data. The output from the first rule statement to generate an action becomes the rule set output for the data row.
- The output from the primary rule set is the primary fact that the rule specification generates about the business data. The primary rule set output represents the result of the business rule analysis of the data row.
- The output from a child rule set is an input to a parent rule set. Every rule set that you add to a rule specification is a child of the primary rule set or a child of another rule set.

  You must use the output from the child rule set in a rule statement in the parent rule set.
- The input properties on a rule set display the inputs that you select in the rule set and the inputs from any child rule set. To view all of the inputs in the rule specification, open the **Manage Global Inputs** dialog box.
- The rule statements in a rule set must generate outputs of the same data type. Otherwise, the rule set might send data of different data types to an input in the parent rule set.

Rule Set Configuration Examples

The number of rule sets that you configure in a rule specification depends on the facts that you need to determine about the input data.

You can configure rule sets in the following ways:

- Configure rule sets that read a single input.
- Configure rule sets that use rule statement conditions in a dependent relationship.
- Configure rule sets that use rule statements in an independent relationship.
- Configure a parent rule set and child rule sets.

Rule Sets that Read a Single Data Input

Configure a rule set to read a single data input when you want to analyze a range of data values in an input column. Define rule statements that apply the same condition logic to the input, and define a different action for the outcome of each condition.

Single Input Example

An airline sells airplane tickets online. The airline sells tickets at different prices for travelers in different age groups. The airline defines a business rule that requires the customer database to identify the age group of each traveler. You configure an input that represents the date-of-birth data in a customer data set. You create a series of rule statements that compare the data values in the input data to the current date.

You configure a rule set with the following rule statements:

```
   IF AGE >= 65 THEN SENIOR
OR IF AGE >= 18 THEN ADULT
OR IF AGE < 18 THEN STUDENT
```

The rule set returns data from the first rule statement that can generate an action. Therefore, the order of the rule statements is relevant. If the rule set reads the second rule statement first, the rule set cannot identify any traveler in the senior age category. When the rule set reads the second rule statement first, the rule set identifies all traveler who are 18 years or older as adults.

Rule Sets with Dependent Conditions

Configure a rule set with rule statement conditions in a dependent relationship when the business rule requires multiple conditions to be true. Define a rule statement that applies a set of related conditions to the inputs, and define a single action for the conditions.

Use an AND operator in the rule statement to join the conditions.

Dependent Conditions Example

A city bus company wants to hire bus drivers. The bus company requires all drivers to pass an eyesight test and to hold a license to drive a bus. The company defines a business rule that requires the employee database to verify the eyesight test status and the license status for each driver. You create inputs that represent the status data for the eyesight test and the license. You configure a rule statement with a condition for each input. You configure the default rule set to return a string when an input to the earlier rule statement is not valid.

You configure a rule set with the following rule statements:

```
      IF EYESIGHT TEST IS YES
      AND DRIVERS_LICENSE IS YES THEN VALID
   OR IF NO RULE STATEMENT IS VALID THEN NOT_VALID
```

Rule Sets with Independent Conditions

Configure a rule set with rule statements in an independent relationship when the business rule requires one of multiple conditions to be true.

Independent Conditions Example

A bank wants to offer credit cards to customers. The bank decides that any customer with a credit balance greater than $5,000 or a salary greater than $50,000 qualifies for a credit card. The bank defines a business rule that requires the customer database to identify the eligible customers. You configure inputs that represent the customer salary and the customer balance. You configure a rule statement to test each input.

You configure a rule set with the following rule statements:

```
      IF CUSTOMER SALARY >= 50000 THEN OFFER_CARD
   OR IF CUSTOMER BALANCE >= 5000 THEN OFFER_CARD
```

The rule set returns data from the first rule statement that can generate an action. However, the order of the rule statements is not relevant in the example. If the data on any input is valid, the input row is valid.

Parent Rule Sets and Child Rule Sets

When a business rule describes a data requirement that depends on other data requirements, configure rule sets in a parent-and-child relationship.

Parent and Child Example

A hardware store sells wood preserver in multiple colors and styles. The store must verify that the product database tracks the inventory of wood preserver in the warehouse. The store defines a business rule that requires the product database to maintains accurate data for the colors, styles, and units of wood preserver. You configure child rule sets with rule statements that verify the data for each type of wood preserver. You configure a parent rule set to verify the output from the child rule sets.

You configure the following child rule sets:

Rule Set Name: Color

```
IF COLOR = (PINE, OAK, CHESTNUT) THEN VALID_COLOR
```

Rule Set Name: Size

```
IF SIZE = (250, 500, 1000) THEN VALID_SIZE
```

Rule Set Name: Style

```
IF STYLE = (GLOSS, MATT) THEN VALID_STYLE
```

You configure the following parent rule set:

Rule Set Name: Wood Preserver

```
IF COLOR = VALID_COLOR
AND SIZE = VALID_SIZE
AND STYLE = VALID_STYLE THEN WOOD PRESERVER IS VALID_PRODUCT
```

To configure a condition to compare an input to a range of values, select the List of Values option in an Input function. You can configure similar parent rule sets and child rule sets for other products in the inventory.

Note: You might also create reference tables that contains lists of the valid colors and styles. Configure the conditions to compare the input values to the reference tables. You can create reference tables in the Analyst tool.

Rule Set Properties

To view the properties on a rule set, select a rule set in the workspace. Use the properties to describe the rule set, to add inputs and rule statements, and to test the rule set.

The rule set properties contain the following views:

General

Displays the rule set name and any description that you add to the rule set. You can update the name and the description.

Enter a name that summarizes the type of data operations that the rule statements perform in the rule set. The rule set name must be unique in the rule specification.

Inputs

Lists the inputs that the current rule set uses. If a rule set is a child to a parent rule set, the parent rule set reads the child rule set name as an input name.

To view all of the inputs in the rule specification or to add an input to the rule specification, open the **Input Management** dialog box.

Rule Logic

Displays the rule statements in the rule set. You can add and update the rule statements.

Test

Displays the rule set inputs in a table that you can use to test the rule set logic. The table includes the inputs from any child rule set below the current rule set.

You can enter sample data values and verify that the rule set generates the results that you expect. To test the complete rule specification, select the primary rule set.

Adding a Rule Set to a Rule Specification

By default, a rule specification contains a primary rule set. Add a rule set under the primary rule set or under another rule set in the rule specification.

1. Open the rule specification in the Design workspace.
2. Select a rule set shape as a parent for the rule set that you create.

   The **Add Rule Set** button appears at the base of the rule set shape.
3. Click the button to add a rule set.

   The Analyst tool adds a rule set shape below the parent rule set.

Cut, Copy, and Paste Operations on Rule Sets

Copy a rule set to reuse the rule statements in the rule set. Move a rule set when you determine that the rule set belongs in another location.

You might copy a rule set to another rule specification to reuse the rule statements that the rule set contains. You might copy a rule set within a rule specification to create rule statements that are similar to the rule statements in the rule set.

You might move a rule set when you determine that the rule set is a dependency for a different business rule requirement.

You copy or move a rule set to a position below another rule set. The rule set that you copy or move becomes a child rule set of the destination rule set. When you copy or move a rule set to another rule specification, you also copy the inputs that the rule set uses. You cannot move a primary rule set. You cannot copy a rule set to the top-level rule specification shape.

Copying and Moving a Rule Set in a Rule Specification

To create multiple similar rule sets in a rule specification, copy the rule set. After you copy the rule set, update the rule statements to reflect the business requirements that the rule set represents.

Move a rule set when you determine that the rule set belongs in another location.

1. Open the rule specification.
2. Click **Edit**.
3. Select a rule set, and right-click it to open the menu.
   - To copy the rule set, select **Copy**.
   - To move the rule set, select **Cut**.
4. Select a parent rule set for the rule set that you copy or move. Right-click the parent rule set to open the menu.

   You can create a copy of a rule set under the current parent rule set.
5. Click **Paste**.
6. The rule set that you copy or move appears below the rule set that you select.

   The Analyst tool adds the string _COPY to the rule set name.
7. Save the rule specification.

Copying and Moving a Rule Set to a Different Rule Specification

If you create multiple rule specifications with similar logic, you can copy the rule sets from one rule specification to another to save time. After you copy the rule set, you might update it to reflect the business requirements of the current rule specification. Or, you might apply the rule set to different inputs.

Move a rule set when you determine that the rule set belongs in another location.

1. Open the rule specification.
2. Click **Edit**.
3. Select a rule set, and right-click it to open the menu.
   - To copy the rule set, select **Copy**.
   - To move the rule set, select **Cut**.
4. On the Analyst tool toolbar, click **Open**.
5. Select Rule Specifications from the list of library assets.
6. Select a destination rule specification for the rule set.

   The rule specification opens in the Design workspace.
7. Click **Edit**.
8. Select a parent rule set for the rule set that you copy or move. Right-click the parent rule set to open the menu.
9. Click **Paste**.
10. The rule set that you copy or move appears below the rule set that you select.
11. Save and close the rule specification.

    If you moved a rule set from another rule specification, save and close both rule specifications.

CHAPTER 4

Rule Statement Configuration

This chapter includes the following topics:

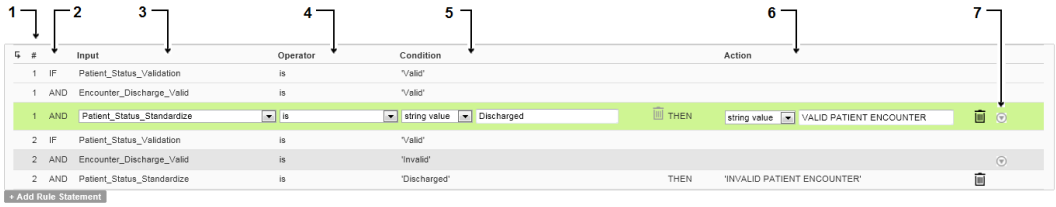
- Rule Statement Configuration Overview, 27
- Conditions, 28
- Operators, 30
- Actions, 31
- Function Expression Descriptions, 34
- Reference Tables, 36
- Rule Statement Menu Options, 37
- Configuring a Rule Statement, 38
- Configuring a Rule Statement that Reads the Result of Another Rule Statement, 39
- Configuring a Rule Statement with Multiple Conditions, 39
- Cut, Copy, and Paste Operations on Rule Statements, 40

Rule Statement Configuration Overview

Rule statements define the data analysis and data transformation operations that a rule specification applies to input data. Rule statements contain inputs, operators, conditions, and actions.

A condition defines the type of analysis that the rule statement performs on the input data. An operator is a mathematical function that determines how the rule statement processes the result of the condition analysis. An action defines the output that the rule statement generates if the input data satisfies the condition.

The following image shows a series of rule statements in the Design workspace:



The rule statements have the following elements:

1. Sequence identifier.
   Indicates the order in which the rule specification runs the rule statements in the rule set.

2. Condition type.
   Indicates the condition or conditions in which the input data can be valid. When you create a rule statement, you create an IF condition. When you add a condition to a rule statement, you create an AND relationship between the conditions in the statement.
3. Input.
   Identifies the data to analyze or update. You can select an input from the current rule set, or you can select the output from a child rule set.
4. Operator.
   Identifies the type of validation operation that the condition applies to the input data.
5. Condition fields.
   Describes a fact that the rule statement verifies about the input data.
6. Action fields.
   Describes the output that the rule statement generates when the input data satisfies the condition. If you configure multiple conditions in the rule statement, the input data must satisfy all conditions to generate an action.
7. Menu options.
   Displays a list of options for the rule statement. You can use the options to link rule statement in the rule set. When you link rule statements, the output from the first rule statement becomes the input to the second rule statement.

Conditions

A condition is a data operation that a rule statement specifies for an input data value. Every rule statement contains at least one condition.

When a condition validates an input data value, the rule specification performs the action in the rule statement. If a condition cannot validate an input data value, the rule specification does not perform the action.

Configure a condition to perform one of the following types of analysis:

Compare the input data values to a single value.

The rule statement compares the values in the input column to the value that you enter.

Compare the input data values to the values on another input.

The rule statement compares the values in the input column to the values on the same row in another input column that you specify.

Compare the input data values to the current date and time.

The rule statement compares the values in the input column to the current date and time on the Data Integration Service host machine.

Search for null values.

The rule statement searches the input column for null or empty strings.

Compare the input data values to a range of values

The rule statement compares the values in the input column to a range of values that you enter.

Compare the input data values to the values in a reference table.

The rule statement compares the values in the input column to the values in a reference table. The rule statement returns a value from the reference table or a value that you enter.

Compare the input data values to the output from a mapplet.

The rule statement compares the values in the input column to the output from a mapplet that you select from the Model repository. You specify one or more inputs for the mapplet. You can select any input in the rule specification, including the current condition input. You can also specify a constant value as a mapplet input.

Apply a function expression to the input data values.

The rule statement applies a function expression that you select to the values in the input column. The Analyst tool stores the list of expressions.

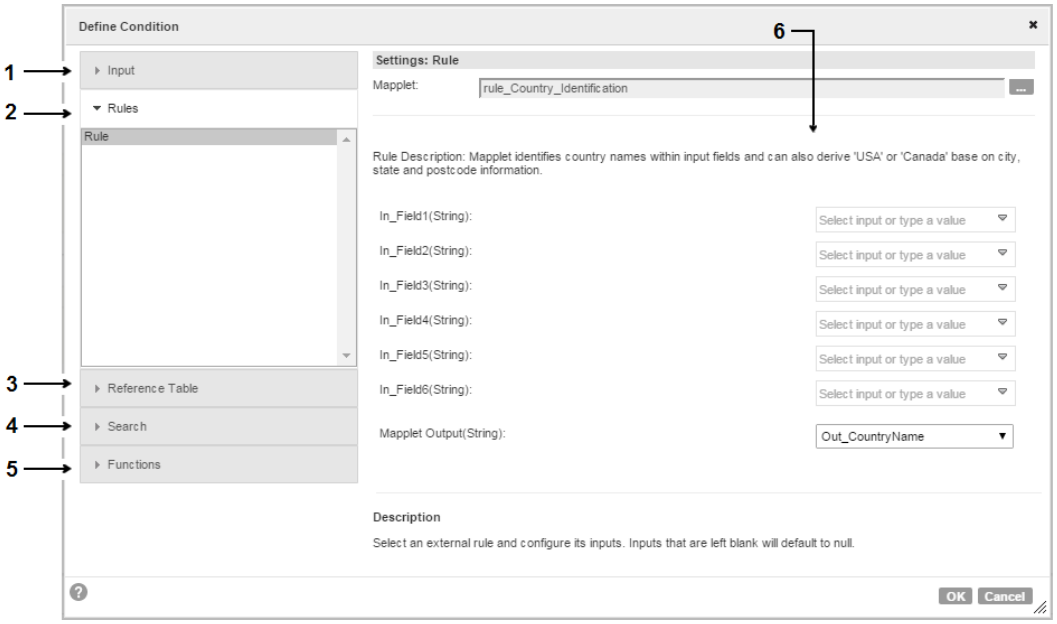
Function Configuration in Conditions

You can configure the data operation in a condition as a function expression. You can also select a function expression from a list of expressions in the rule specification. Use the function options in the condition to configure or select the expression.

You can use the function options as an alternative to the inline condition options that the rule statement displays in the workspace. You can also use the function options to configure a condition that reads multiple inputs or that reads an asset from the Model repository.

Use the **Define Condition** dialog box to configure or select a function expression. To open the dialog box, select *function* as the condition type in the rule statement.

The following image shows the **Define Condition** dialog box:



The dialog box contains the following options:

1. Input
   Use the Input options to compare an input data value to other data values in the following ways:
   - Compare the rule statement input with an input from another column on the same row.

- Compare the rule statement input with a value that you enter.
- Compare the rule statement input with a range of values that you enter.

2. Rules
   Use the Rules options to compare an input data value to the output from a mapplet. Use the options to select the mapplet from the Model repository. The Library workspace identifies the mapplets that you can select as rules.
3. Reference Table
   Use the Reference Table options to compare an input data value with the values in a reference table. Use the options to select the reference table from the Model repository. The Library workspace lists the reference tables that you can select.
4. Search
   Use the Search options to find an input data value in another input column. The rule statement searches all of the values in the column for the current rule statement input value.
5. Function
   Use the Function options to apply a function expression to the input data. Select an expression from the list in the dialog box.
6. Configurable options for the operation
   Use the options to configure the condition.

Operators

An operator is a mathematical function that determines the outcome of a condition.

You can select one of the following operators:

is

Verifies that the input data matches the data that the condition specifies.

is not

Verifies that the input data does not match the data that the condition specifies.

is within

Verifies that the input data matches a value in a reference table or in a list of values that you enter.

is not within

Verifies that the input data does not match a value in a reference table or in a list of values that you enter.

contains

Verifies that the input data contains a sequence of characters that you specify.

is less than

Verifies that the input data value is less than the value that the condition specifies. Select the operator with numeric input data.

is less than or equal to

Verifies that the input data value is less than or equal to the value that the condition specifies. Select the operator with numeric input data.

is greater than

Verifies that the input data value is greater than the value that the condition specifies. Select the operator with numeric input data.

is greater than or equal to

Verifies that the input data value is greater than or equal to the value that the condition specifies. Select the operator with numeric input data.

Actions

An action specifies the output from a rule statement. An action generates an output value when the conditions in the rule statement validate an input value.

Configure an action to generate an output in one of the following ways:

Return a value that you enter.

Configure an action to return a data value that you specify.

For example, you might configure an action to return the word VALID when the input data matches a reference data value.

Return a value from an input column.

Configure an action to return a data value from an input in the rule specification. The action reads the data value on the corresponding row of the input that you select.

For example, a business might define a rule that requires all product records to use Quick Response (QR) codes in place of bar codes. You configure an action to replace an input column of bar code values with a corresponding column of QR code values.

Return a value from a reference table.

Configure an action to return a data value from a reference table. The action compares the input data to the values in the reference table that you specify. If the action finds the input data in the reference table, the action returns the reference data value that corresponds to the input data.

For example, you might configure an action to compare first name data to a reference table that identifies the gender of the first name. The action might read the word "JOHN" from the input data and return the word "MALE" from the reference table.

Return a value that you enter when an input data value matches a reference table value.

Configure an action to return a data value from a reference table. The action compares the input data to the values in the reference table that you specify. If the action finds the input data in the reference table, the action returns the string that you entered. if the action does not find the input data in the reference table, the action returns the input value.

For example, you might configure an action to compare employee code data to a reference table that contains the valid employee codes in an organization. You define a function that returns the word "VALID" when an employee code matches a value in the reference table. If the function does not find an employee code in the reference table, the action returns the employee code from the input column.

Note: The action returns a data value for all records that satisfy the condition in the rule statement.

Return the current date and time.

Configure an action to return the system date and time to nanosecond precision from the Data Integration Service host machine. The rule statement returns the date and time values in a date/time data type. You must specify the format for the date and time.

Remove data values from the input data.

Configure an action to remove a data value that you specify.

For example, the input data might include a product name that the business no longer uses. Configure the action to return a version of the input data that omits the product name.

Remove reference data values from the data.

Configure an action to remove reference data values from an input.

The action compares the input data to the values in a reference table that you specify. If the action finds an input data value in the reference table, the action returns a version of the input without the data value.

For example, you might configure an action to compare an input string to a reference table of salutation terms. The action might read the input string "MR JOHN SMITH" and return the string "JOHN SMITH" as the output.

Remove character spaces from the input data.

Configure an action to delete character spaces from the input data. Remove character spaces when the input data includes redundant character spaces.

For example, you might configure an action to remove character spaces from the following telephone number:

```
212 555 1234
```

Use a function expression to determine the output value.

Configure an action to apply a function expression to the input data.

For example, you might configure the action to return the character length of the input data.

Merge values from multiple inputs.

Configure an action to merge data from two or more inputs.

For example, you might configure an action to merge a column of first name data and a column of second name data. The action returns a single field that contains the first name and second name.

Change the character case of the input data.

Configure an action to convert the character case of an input that you select.

For example, a rule statement might read a data set that stores data values in sentence case. Configure an action to return the data in uppercase.

Apply a mapplet to the input data.

Configure an action to apply a mapplet to the input data and to return the mapplet output. Select the input from the rule specification, or enter a constant value as the mapplet input. Select the mapplet from the Model repository.

A mapplet might represent a set of operations that the organization approves for use in for business rules. Or, a mapplet might contain transformation logic that you cannot define in the other options in the rule statement.

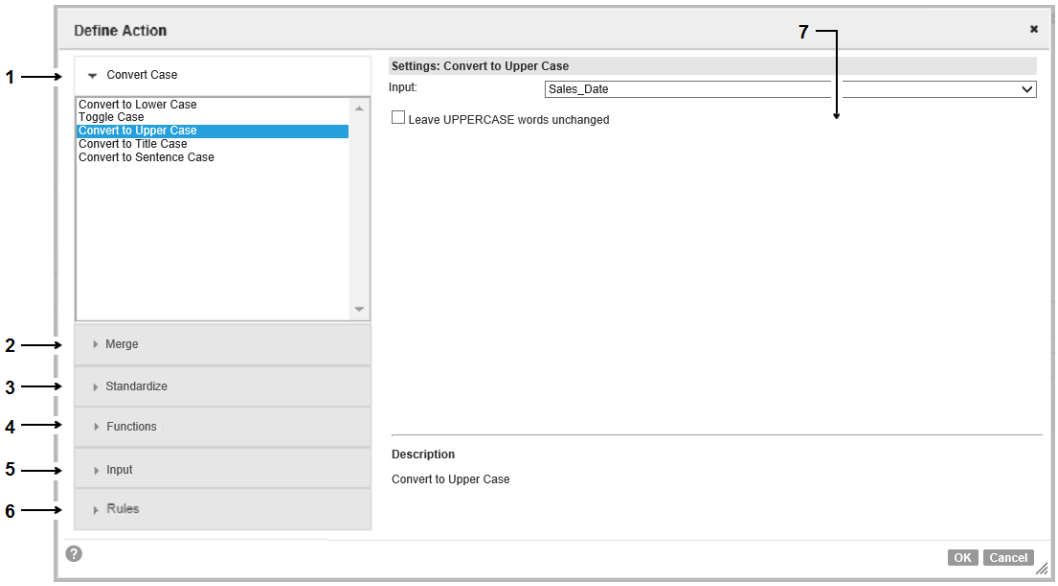
Function Configuration in Actions

You can configure the data operation in an action as a function expression. You can also select a function expression from a list of expressions in the rule specification. Use the function options in the action to configure or select the expression.

You can use the function options as an alternative to the inline action options that the rule statement displays in the workspace. You can also use the function options to configure an action that reads multiple inputs or that reads an asset from the Model repository.

To configure an expression, select the function option from the action menu. When you select the field, you open the **Define Action** dialog box.

The following image shows the **Define Action** dialog box:



The dialog box contains the following options:

1. Convert Case
   Use the Convert Case options to change the character case of the input data values.
2. Merge
   Use the Merge options to return data values from multiple inputs as a single output.
3. Standardization
   Use the Standardization options to update a copy of the input data value and to return the copy. You can configure a standardization function to remove characters, to return a value from a reference table, or to return a value that you specify.
4. Function
   Use the Function options to apply a function expression to the input that you select and to return the output of the expression.
5. Input
   Use the Input options to return a data value. You can specify a constant value, or you can select an input. If you select an input, the action returns the value in the input column on the same row as the condition input.
6. Rules
   Use the Rules options to return the output of a mapplet that you select.

7. Configurable options for the operation
   Use the options to configure the action.

Function Expression Descriptions

A function expression is a logical element in a condition or an action. A function expression reads the input data that you specify, performs a calculation on the input data, and returns a result.

When you add a function expression to a condition, the function expression performs the calculation and returns the result to the condition. The condition uses the rule statement operator to determine the next step. When you add a function expression to an action, the action performs the calculation and returns the result as the rule statement output.

To add a function expression to a condition, use the Function options in the **Define Condition** dialog box. To add a function expression to an action, use the Function options in the **Define Action** dialog box.

You can add the following function expressions to a condition or an action:

Add to Date

Reads a date value and adds a numeric value to an element in the date. Enter the value to add to the element, or select an input that contains the value to add. Select the function expression in a condition or an action.

Choose

Reads input data that contains multiple values and selects the value at the position that you specify. Select the function expression in a condition or an action.

For example, the following string contains values at four positions:

```
Q1, Q2, Q3, Q4
```

You can enter a position, or you can select an input to specify a position.

Convert to Date

Converts a date value in an input string to a date/time data type. The function expression returns the month, day, and year. You select the date format when you configure the function expression. Select the function expression in an action.

Concatenate

Reads two inputs that you select and concatenates the values from the inputs. Select the function expression in a condition or an action.

Date Difference

Reads date values from two inputs that you select and calculates the numeric difference between the dates. You select the date element that the function expression uses to measure the difference. Enter the date element, or select an input that contains the date element. Select the function expression in a condition or an action.

Date Part

Reads a date from an input and identifies the numeric value of a date element that you specify. Enter the date element, or select an input that contains the date element. Select the function expression in a condition or an action.

Date Time

Uses multiple values to specify a date and time. Enter a value for each date element, or select an input to add a value to each element. Select the function expression in a condition or an action.

Greatest

Reads two or more values and identifies the highest value. Enter the values, or select rule set inputs to add each value. You can specify numeric values, alphabetic values, or date values. Select the function expression in a condition or an action.

When you specify numeric values, the function expression identifies the highest number.

When you specify alphabetic values, the function expression identifies the value that is last in alphabetic order.

When you specify date values, the function expression identifies the latest date.

Last Day

Reads a date value from an input and identifies the last day in the month that the date specifies. The function expression identifies the last day as a numeric value. Select the function expression in a condition or an action.

Least

Reads two or more values and identifies the lowest value. Enter the values, or select rule set inputs to add each value. You can specify numeric values, alphabetic values, or date values. Select the function expression in a condition or an action.

When you specify numeric values, the function expression identifies the lowest number.

When you specify alphabetic values, the function expression identifies the value that is first in alphabetic order.

When you specify date values, the function expression identifies the earliest date.

Length

Reads an input value and calculates the number of characters in the value. Select the function expression in a condition or an action.

Lower

Reads an input string and returns the lowercase version of the characters in the string. Select the function expression in a condition.

Null

Reads an input value and determines if the value is null. Select the function expression in a condition or an action.

Replace Character

Reads an input value and replaces a sequence of characters in the value with a sequence that you specify. Select the function expression in a condition.

Replace String

Reads an input value and replaces the value with a value that you specify. Select the function expression in a condition.

Reverse

Reads an input value and reverses the order of the characters in the value. Select the function expression in a condition or an action.

Truncate

Reads an input value and truncates the value to the number of places that you specify. The function expression reads a numeric value. Enter the value, or select an input to add the value. Select the function expression in a condition or an action.

Upper

Reads an input string and returns the uppercase version of the characters in the string. Select the function expression in a condition.

Date Elements in Functions

When you add a date element to a function, enter the function in the format that the function recognizes.

The following table describes the date formats to use:

Date Element	Format
Year	yyyy
Month	MM
Day	dd
Hour	hh
Minute	mm
Second	ss
Millisecond	SSS

Reference Tables

You can add a reference table to a condition or an action. Add a reference table to a condition to verify that the input data contains the data values that you expect. Add a reference table to an action to return a value from the table that corresponds to the input data value.

A reference table contains the standard versions of a set of business values. A reference table might contain public terms, such as telephone area codes or address abbreviations. Or, a reference table might contain values that are specific to an organization, such as employee codes or product codes. You can create reference tables in the Analyst tool, or you can ask a developer to create a reference table. You select a reference table from a project in the Model repository when you configure the condition or the action.

A reference table contains two or more columns. Each row in the table represents a single business value. A row might contain two or more identical values, or it might contain alternative versions of the same value, including incorrect values.

If you add the reference table to a condition, the condition compares the input data to the values in the first column in the reference table. If the condition finds an input value in the reference data, the condition applies the operator to the input data to determine the action to take.

If you add the reference table to an action, the action compares the input data to the values in the second or additional columns. If the action finds an input data value in the reference table, the action returns the value on the same row in the first column.

Reference Table Example

The following table contains sample reference table data:

Name1	Name2
Mumbai	Bombay
Chennai	Madras
Bengaluru	Bangalore
Puducherry	Pondicherry
Kadapa	Cuddapah
Kochi	Cochin
Kolkata	Calcutta

The Name1 column contains the current names of cities in India. The Name2 column contains the previous version of each name. Add the reference table to a condition to verify that the input data contains the current city names. Add the reference table to an action to find the older city names and replace them with the current city names.

Rule Statement Menu Options

You can open a set of menu options for each rule statement. Use the menu options to organize the rule statements in the rule set.

The menu has the following options:

Add Condition

Add a condition to the current rule statement. When you configure multiple conditions in a rule statement, you create an AND relationship between the conditions.

Add Rule Statement

Add an empty rule statement to the rule set.

Add Rule Statement Using Result

Add a rule statement to the rule set below the current rule statement. The rule statement that you add uses the output from the current rule statement as an input.

Copy Rule Statement

Copy the current rule statement. You can paste the rule statement to the current rule set or to another rule set.

Cut Rule Statement

Remove the current rule statement from the rule set. You can paste the rule statement to another rule set.

Move Down

Reorder the rule statements so that the current rule statement changes places with the rule statement below it.

Move to Bottom

Move the current rule statement to the bottom of the rule set.

Move to Top

Move the current rule statement to the top of the rule set.

Move Up

Reorder the rule statements so that the current rule statement changes places with the rule statement above it.

Paste Rule Statement

Paste a rule statement that you copied or moved in an earlier action.

Configuring a Rule Statement

Before you configure the rule statement, verify that the rule specification contains the inputs that the rule statement requires.

1. Open the rule specification.
2. Select the rule set to contain the rule statement.
3. Click **Rule Logic**.
4. Click **Add Rule Statement**.
5. Add an input to the rule statement. The condition reads the input that you add.
6. Select an operator. The operator specifies the type of comparison operation that the condition performs.
7. Configure the condition.

   Identify one or more data values that the condition can compare to the rule statement input. Or, configure a data operation to apply to the input.
8. Configure the action to perform when the condition validates the input data.

   Identify the input data or a constant data value that the action can return as the rule statement output. Or, configure a data operation that defines the output. Use the function options to configure the operation.
9. Save the rule specification.

Configuring a Rule Statement that Reads the Result of Another Rule Statement

You can create a chain of rule statements in which each statement reads the result of the statement above it. Use the rule statements to derive information from the relationship between multiple inputs.

1. Open the rule specification, and click **Edit**.
2. Select the rule set.
3. Click **Rule Logic**.
4. Select the rule statement that generates the result that another rule statement must read.

   Or, configure the rule statement.
5. From the rule statement menu, select **Add Rule Statement Using Result**.

   The Analyst tool adds an empty rule statement under the rule statement that you select.
6. Configure and save the rule statement.

Note: Data passes through a chain of rule statements when all of the logical operations in the statements are valid for the inputs that you select. If a rule statement in the chain does not generate a result for a data row, the chain performs no further analysis on the row.

Configuring a Rule Statement with Multiple Conditions

You can configure multiple conditions in a single rule statement. When you use multiple conditions, you create an AND relationship between the conditions.

You configure a single action in the rule statement.

1. Open the rule specification.
2. Select the rule set to contain the rule statement.
3. Click **Rule Logic**.
4. Click **Add Rule Statement**.
5. Add an input to the rule statement. The condition reads the input that you add.
6. Select an operator. The operator specifies the type of comparison operation that the condition performs.
7. Configure the condition.

   Identify one or more data values that the condition can compare to the rule statement input. Or, configure a data operation to apply to the input.
8. From the rule statement menu, select **Add Condition**.

   The Analyst tool adds a condition to the workspace.
9. Configure the condition.

   You can add multiple conditions.
10. Configure the action for the rule statement.

    The Analyst tool adds the action after the final condition in the rule statement.
11. Save the rule specification.

Cut, Copy, and Paste Operations on Rule Statements

Copy a rule statement to reuse the rule statement logic in the same rule set or in another rule set. Move a rule statement that belongs in another rule set. You can copy or move a rule statement to a rule set in the same rule specification or in another rule specification.

You might copy a rule statement to another rule specification to reuse the rule statement. You might copy a rule statement within a rule specification to create a similar rule statement.

You might move a rule statement when you determine that the rule statement applies to a different business rule requirement.

When you copy or move a rule statement to another rule specification, you also copy the inputs that the rule statement uses.

Copying and Moving a Rule Statement in a Rule Specification

To create multiple similar rule statements in a rule specification, copy the rule statement. You can copy and paste a rule statement within the same rule set or to another rule set. After you copy the rule statement, update it to reflect the business rule logic.

Move a rule statement when you determine that the rule statement belongs in another rule set.

1. Open the rule specification.
2. Click **Edit**.
3. Select the rule set that contains the rule statement to copy or move.
4. Select **Rule Logic**.
5. Select the rule statement. If the rule statement contains more than one condition, select the first line of the rule statement.
6. Open the context menu.
   - To copy the rule statement, select **Copy Rule Statement**.
   - To move the rule set, select **Cut Rule Statement**.

   Note: Find the context menu option beside the Action options.
7. Select a parent rule set for the rule statement.

   You can create a copy of a rule statement in the current rule set or in another rule set.
8. Select a rule statement in the rule set.
9. Open the rule statement context menu, and select **Paste Rule Statement**.

   The rule statement that you copied or moved appears in the rule set.
10. Verify the position of the rule statement in the rule set.

    If necessary, move the rule statement up or down in the list of list of rule statements.
11. Save the rule specification.

Copying and Moving a Rule Statement to a Different Rule Specification

When you create multiple rule specifications with similar logic, you can copy the rule statements from one rule specification to another to save time. After you copy the rule statement, you might update it to reflect the

business requirements of the current rule specification. Or, you might apply the rule statement to different inputs.

Move a rule specification when you determine that the rule set belongs in another location.

1. Open the rule specification.
2. Click **Edit**.
3. Select the rule set that contains the rule statement to copy or move.
4. Select **Rule Logic**.
5. Select the rule statement. If the rule statement contains more than one condition, select the first line of the rule statement.
6. Open the context menu.
   - To copy the rule statement, select **Copy Rule Statement**.
   - To move the rule set, select **Cut Rule Statement**.

   Note: Find the context menu option beside the Action options.
7. On the Analyst tool toolbar, click **Open**.
8. Select Rule Specifications from the list of library assets.
9. Select the name of the destination rule specification for the rule set.

   The rule specification opens in the Design workspace.
10. In the open rule specification, click **Edit**.
11. Select a rule set, and select **Rule Logic**.
12. Select a rule statement in the rule set.
13. Open the rule statement context menu, and select **Paste Rule Statement**.

    The rule statement that you copied or moved appears in the rule set.
14. Verify the position of the rule statement in the rule set.

    If necessary, move the rule statement up or down in the list of list of rule statements.
15. Save and close the rule specification.

    If you moved a rule statement from another rule specification, save and close both rule specifications.

CHAPTER 5

Common Types of Rule Statements

This chapter includes the following topics:

- Common Types of Rule Statements Overview, 42
- Verifying the Accuracy of Business Data , 44
- Identifying Obsolete Values in Business Data, 44
- Standardizing Business Data Values, 45
- Improving the Usability of Business Data , 46
- Discovering Information About Business Users, 47
- Using Data Values to Verify a Business Policy, 48
- Updating Data Values to Meet Business Standards, 49
- Verifying the Application of a Business Rule, 49
- Validating Address Records, 50
- Using a List of Values to Find Records in a Data Set, 52
- Identifying Data Values that Contain Key Words or Character Strings, 53

Common Types of Rule Statements Overview

You can define all or part of a business rule in a rule statement. Define a rule statement for each fact that a business rule must discover. The type of rule statement that you define depends on the business rule requirements.

You can use rule statements to verify and update the following business facts:

Verify the accuracy of business data

Define a rule statement that compares input data to a reference table.

Use the rule statement to verify that the business data is accurate. For example, define a rule statement to verify that a product description table uses the current product codes.

Find obsolete business data

Define a rule statement to find values that are no longer valid or relevant to the business.

Use the rule statement to return a user-defined value that identifies any row that contains an obsolete term.

Standardize business data values

Define a rule statement that searches for values that you specify and replaces them with other values.

Use the rule statement to verify that the business data follows a standard format. For example, define a rule statement to verify that financial data does not use currency symbols. The rule statement replaces the currency symbols with agreed abbreviations for the currency names.

Improve the usability of business data

Define a rule statement that removes redundant characters and redundant strings.

Use the rule statement to ensure that users and software applications can read the business data correctly. For example, define a rule statement to delete character spaces at the start and end of a data field.

Use data values to discover information about business users

Define a rule statement that correlates information from different inputs. Use different conditions to analyze the inputs, and connect the conditions with AND logic. For example, define conditions to link different facts about customers and to derive information that can improve customer service.

Use data values to verify a business policy

Define a rule statement that verifies that the business follows a stated policy.

Use the rule statement to define a function expression that represents the policy.

Update data values to match business standards

Define a rule statement that adds characters to data values so that the values meet the current business standards.

Use the rule statement to define an expression that updates the values.

Verify that you applied a business rule to a data set

Define a rule statement that adds the current date and time to a data set.

Use the rule statement to add a time stamp to the data set that indicates when the mapping ran.

Verify postal address data

Define a rule statement that applies an address validation mapplet to the input data. Define another rule statement that evaluates the output from the address validation mapplet.

Use a list of values to find data in a column

Define a rule statement that looks for the values that you enter in a column of data.

Use the rule statement to define a list of data values when the Model repository does not contain a reference table of the values.

Identify Data Values that Contain Key Words or Character Strings

Define a rule statement that finds a data value or a character string in fields that contain multiple values or longer character strings.

Use the rule statement to find data values that contain similar information.

Verifying the Accuracy of Business Data

To verify the accuracy of a column of business data, configure a rule statement that compares the column to a reference table. For example, you might configure a rule statement to verify that a data column contains the current product codes for the organization.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.
5. Select an operator to validate the results of the condition analysis.
   - To identify input values that match the reference table values, select the following operator:

     ```
     is within
     ```
6. Select the type of condition to apply to the input.
   - To use a reference table in a condition, select the following condition type:

     ```
     function
     ```
7. Configure the condition to apply a reference table to the input data.
   - Open the **Define Condition** dialog box.
   - Select **Reference Table**.
   - Browse to the reference table that contains the standard versions of the business data.
     For example, select a reference table that contains a set of product codes.
   - Select or clear the option to perform a case-sensitive search of the reference table.
   - Click **OK**.
8. Select the type of action to apply to the data that satisfies the condition logic.
   - To return a string value, select the following action type:

     ```
     string value
     ```
9. Enter the value that the action returns. For example, enter VALID.
10. Save the rule specification.

Identifying Obsolete Values in Business Data

To find values that are no longer valid in a data set, configure a rule statement to read a reference table that contains known data errors. For example, a national organization updates the codes that identify branch offices. The organization changes the code structure from eight digits to ten digits. You might configure a rule statement to read a reference table that contains the obsolete eight-digit codes.

If the rule statement finds an obsolete branch code in the input data, it writes a value that you specify as output. If the rule statement does not find an obsolete branch code, the action writes the input value as output. The action writes an output for each input row that satisfies the condition in the rule statement.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.

3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the branch code data.
5. Select an operator to validate the results of the condition analysis.
   - To specify that the input data must match the data in the condition, select the following operator:

     `is`
6. Select the type of condition to apply to the input.
   - To compare the input data to the data from another input, select the following condition type:

     `input`
7. Select the input that you selected in step 4.

   The rule statement compares the input to itself. Therefore, the rule statement passes all of the input data values from the condition to the action.
8. Configure the action to search the input data for obsolete branch codes.
   - Open the **Define Action** dialog box.
   - Select **Standardize**, and select **Replace Reference Table Matches with Custom Strings**.
   - Select the input that you specified in the condition.
   - Select the reference table from the Model repository.
   - Enter the data value that the action returns when an input data value matches a reference table value. For example, enter OBSOLETE_CODE.
9. Click **OK**.
10. Save the rule specification.

Standardizing Business Data Values

To standardize data values, configure a rule statement that replaces the values with the value that you require. For example, you might configure a rule statement to replace any currency symbol in a data column with the agreed abbreviation for the currency name.

To standardize multiple values, you can create additional rule statements. Alternatively, use a reference table that identifies the values.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.
5. Select an operator to validate the results of the condition analysis.
   - To specify that the input data must match the data that you enter, select the following operator:

     `is`
6. Select the type of condition to apply to the input.

- To compare the input data to a string value, select the following condition type:

  ```
  string value
  ```

7. Enter the value to compare to the input data. For example, enter $.
8. Select the type of action to apply to the data that satisfies the condition logic.
   - To return a string value, select the following action type:

     ```
     string value
     ```

9. Enter the value that the action returns. For example, enter USD.
10. Save the rule specification.
11. Repeat the steps for other currency symbols that you want to standardize.

    For example, standardize £ to GBP.

Improving the Usability of Business Data

To improve the usability of data values, you can delete character spaces at the start and at the end of a data field. The redundant spaces introduce errors when you sort or compare data values that contain the same information.

Configure a rule statement that deletes the redundant spaces. To delete longer character strings, create a rule statement that uses the "Replace Character" function.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.
5. Select an operator to validate the results of the condition analysis.
   - To specify that the condition does not apply to null data, select the following operator:

     ```
     is not
     ```

6. Select the type of condition to apply to the input.
   - To verify that the input values are not null, select the following condition type:

     ```
     null value
     ```

7. Select the type of action to apply to the data that satisfies the condition logic.
   - To standardize the data, select the following action type:

     ```
     function
     ```

8. Configure the action to remove the redundant spaces from the input data.
   - Open the **Define Action** dialog box.
   - Select **Standardize**, and select **Remove Leading and Trailing Spaces**.
   - Select the input that you specified in the condition.
   - Click **OK**.
9. Save the rule specification.

Discovering Information About Business Users

To discover information about business users, configure a rule statement that measures a fact about the users. When you combine multiple conditions to measure related facts, you discover additional information.

For example, you might configure a rule statement for a financial institution to identify customers who hold credit balances. You might also configure a rule statement to identify customers who earn a high salary.

Use the AND feature to combine the rule statements.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.

     Note: The rule statement analyzes salary data.
5. Select an operator to validate the results of the condition analysis.
   - To identify input values that indicate a minimum salary of $100,000, select the following operator:

     `is greater than or equal to`
6. Select the type of condition to apply to the input.
   - To compare the input data to an integer, select the following condition type:

     `integer value`
7. Enter the value to compare to the input data. For example, enter 100000.
8. Select **Add Condition** in the rule statement menu.

   The rule specification adds a condition under the current condition and creates an AND relationship between the conditions.
9. Select an input for the condition that you added.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.

     Note: The rule statement analyzes the customer account balances.
10. Select an operator to validate the results of the condition analysis.
    - To identify input values that indicate a minimum balance of $10,000, select the following operator:

      `is greater than or equal to`
11. Select the type of condition to apply to the input.
    - To compare the input data to an integer, select the following condition type:

      `integer value`
12. Enter the value to compare to the input data. For example, enter 10000.
13. Select the type of action to apply to the data rows that meet all of the conditions in the rule statement.
    - To return a string value, select the following action type:

      `string value`
14. Enter the value that the action returns. For example, enter HIGH.
15. Save the rule specification.

Using Data Values to Verify a Business Policy

To verify that the business follows a stated policy, define a rule statement that applies an expression to key data columns.

For example, you might configure a rule statement that validates mortgage loan applications for a mortgage broker. The mortgage broker refuses any loan application that requires the applicant to repay more than twenty-five percent of their salary each month. The rule statement tests the monthly repayment amount against an expression that calculates the salary percentage.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.

     Note: The rule statement analyzes the monthly repayment amount on the mortgage loan.
5. Select an operator to validate the results of the condition analysis.
   - To identify input values that are greater than twenty-five percent of another value, select the following operator:

     ```
     is greater than
     ```
6. Select the type of condition to apply to the input.
   - To define an expression to test the input data, select the following condition type:

     ```
     function
     ```
7. Configure the condition to apply an expression to the input data. The expression uses an input that represents the monthly salary of the loan applicant.
   - Open the **Define Condition** dialog box, and select **Input**.
   - Select the float or the integer data type.
   - Select the input column that represents the monthly salary data.
   - Add an empty value under the current value.
   - Select the division sign to indicate that the expression divides the first value by the second value.
   - Click **User Defined Value**, and enter the number four.
   - Click **OK**.
8. Select the type of action to apply to the data that satisfies the condition logic.
   - To return a string value, select the following action type:

     ```
     string value
     ```
9. Enter the value that the action returns. For example, enter DENY_LOAN.
10. Save the rule specification.

Updating Data Values to Meet Business Standards

To update the values in a column without loss of data in the column, define a rule statement that applies an expression to the column.

For example, you might configure a rule statement that updates a column of product codes. The product owners decide to append a date prefix to the current codes. The rule statement uses an expression to update the codes.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that represents the business data.

     Note: The rule statement reads string data.
5. Select an operator to verify that the input values are not null.
   - Select the following operator:

     ```
     IS NOT
     ```
6. Select the type of condition to apply to the input.
   - To verify that the input values are not null, select the following condition type:

     ```
     null value
     ```
7. Select the type of action to apply to the data that satisfies the condition logic in both rule statements.
   - To apply an expression to the data, select the following action type:

     ```
     function
     ```
8. Configure the action to apply an expression to the input data. The expression uses an input that represents the current product codes and a string value to append to each code.
   - Open the **Define Action** dialog box, and select **Input**.
   - Select String as the data type.
   - Add an empty value under the current value.
   - Enter the value to append to the current product codes as the first value.
   - Select the input that represents the product codes as the second value.
   - Click **OK**.
9. Save the rule specification.

Verifying the Application of a Business Rule

To demonstrate that you applied a business rule to a data set, you might add a date and time stamp to each record. The date and time stamp indicate the time at which the Data Integration Service ran a mapping to validate the business rule on the data set.

For example, you might define a business rule that states that primary key columns in a data set cannot contain null values. You define a parallel business rule that sets a monthly schedule to review the primary key

column data. You configure a rule statement that analyzes the primary key columns and that adds a date and time stamp to indicate the time of the analysis.

Add a condition for each primary key column. Use the AND feature to combine the conditions.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the primary key data. For example, create a string input that can contain Social Security numbers.
5. Select an operator to validate the results of the condition analysis.
   - To specify that the condition does not apply to null data, select the following operator:

     ```
     is not
     ```
6. Select the type of condition to apply to the input.
   - To verify that the input values are not null, select the following condition type:

     ```
     null value
     ```
7. Select **Add Condition** in the rule statement menu.

   The rule specification adds a condition under the current condition and creates an AND relationship between the conditions.
8. Select an input for the rule statement that you added.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data. For example, create a string input that can contain account numbers.
9. Repeat steps four through eight to configure a condition for any additional primary key column in the data set.
10. Select the type of action to apply to the data rows that meet all of the conditions in the rule statement.
    - When you generate a mapplet from the rule specification, a developer adds the mapplet to a mapping and runs the mapping on the data set.

      To return the date and time at which the mapping runs, select the following action type:

      ```
      current time stamp
      ```
11. Save the rule specification.

Validating Address Records

To verify that a set of address records are valid for postal delivery, add an address validation mapplet to the action in a rule statement. Link the rule statement to another rule statement that analyzes the mapplet output.

For example, you might select a mapplet that returns a delivery status code for each address in a data set. Configure a condition to verify that the address data column is not null. Configure an action to apply the mapplet to the data. Configure an additional rule statement to evaluate the status codes that the mapplet generates.

To identify the mapplet to use, discuss the business rule requirements with a Developer tool user.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent a line of address data. Set a maximum length for the input that can accommodate the address line.
5. Select an operator to validate the results of the condition analysis.
   - To specify that the condition does not apply to null data, select the following operator:

     ```
     is not
     ```

     Select the following condition type:

     ```
     null value
     ```
6. Select the type of action to apply to the data that satisfies the condition logic.
   - To use a mapplet, select the following action type:

     ```
     function
     ```
7. Configure the action to apply the mapplet to the input data.
   - Open the **Define Action** dialog box.
   - Select **Rules**, and then select **Rule**.
   - Browse to the address validation mapplet.
   - Connect the address line data input to the mapplet. Optionally, connect other rule specification inputs to the mapplet. You can also enter a constant value as a mapplet input.
   - Select a mapplet output as the result of the action. For example, select the status code data that the business rule specifies.
   - Click **OK**.

   Note: The mapplet that you select can run with a single input. If you do not connect an input on the mapplet, the input uses null data. The mapplet might contain multiple outputs. You select a single output as the result of the action. The mapplet does not generate data on the outputs that you do not select. Ask the mapplet developer to identify the inputs and the outputs on the mapplet.
8. Select **Add Rule Statement Using Result** from the rule statement menu. The Analyst tool adds a rule statement under the current rule statement.

   The output from the current rule statement becomes the input to the second rule statement. An arrow connects the rule statements, and the rule statements use the same identifier.
9. Select an operator to validate the results of the condition analysis.
   - To specify that the input data must match the data that you enter, select the following operator:

     ```
     is
     ```
10. Select an operator to validate the results of the condition analysis.
    - To compare the result of the previous rule statement to a list of status codes, select the following operator:

      ```
      is within
      ```
11. Select the type of condition to apply to the input.
    - To specify a list of values, select the following operator:

      ```
      function
      ```

12. Configure the condition to compare the mapplet output to a range of status codes. The status code uses an alphabetic character to indicate the type of validation that the mapplet performed. The code uses a digit to indicate the accuracy of the address.
    - Open the **Define Condition** dialog box.
    - Select **Input**, and select **List of Values**.
    - Set the data type to integer.
    - Enter the following values:

      ```
      V4, V3, C4, C3
      ```

      Enter each value on a separate line.
    - Click **OK**.
13. Select the type of action to apply to the data that satisfies the condition logic in both rule statements.
    - To return a string value, select the following action type:

      ```
      string value
      ```
14. Enter the value that the action returns. For example, enter VALID_ADDRESS.
15. Save the rule specification.

Using a List of Values to Find Records in a Data Set

To determine if any value in a list of values appears in a column of business data, add the values on the list to a condition. The rule statement compares the values in the input that you select to the values in the list.

Add a list of values to a condition in the following scenario:

- The business rule requirement that the condition represents applies to multiple values.
- The organization does not maintain a reference table of the values.

For example, you might work for a manufacturing company that must recall a small number of products because of a product defect. You know the serial numbers of the products that you must recall. You configure a rule statement with an input that represents the column of serial number data. You add the list of serial numbers to recall to the rule statement condition.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
    - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.
5. Select an operator to validate the results of the condition analysis.
    - To identify input values that match the reference table values, select the following operator:

      ```
      is within
      ```
6. Select the type of condition to apply to the input.
    - To enter a list of values, select the following condition type:

      ```
      function
      ```
7. Configure the condition to apply a reference table to the input data.

- Open the **Define Condition** dialog box.
- Select **Input**, and select **List of Values**.
- Enter a list of values. By default, the dialog box displays a single data field. Add a data field for each value on the list.
- Click **OK**.

8. Select the type of action to apply to the data that satisfies the condition logic.
   - To return a string value, select the following action type:

     `string value`
9. Enter the value that the action returns. For example, enter RECALL.
10. Save the rule specification.

Identifying Data Values that Contain Key Words or Character Strings

To find a word or a character string in a column of input data, configure a condition to use the "contains" operator. You might use the operator to find a single data value when each field in the input contains multiple values. Or, you might use the operator to find input data values that contain a character string within a longer string.

For example, you might work for an organization that stores inventory items in a warehouse. The organization must manage the warehouse space and ensure that inventory items stay in the warehouse for as little time as possible. You want to measure the number of items in inventory for a range of products. You configure a single condition to identify the products in a product table. The product names are West Wood, West Star, and West Land. You design a condition to find product names that contain the string "West."

The rule statement that you configure might represent a single element in a business rule. For example, the business rule might describe a policy to reduce the sale price of products that spend a long time in inventory. The current rule statement generates an output that a rule statement in a parent rule set can read as an input.

1. Select a rule set in a rule specification.
2. In the rule set properties, click **Rule Logic**.
3. Click **Add Rule Statement**.
4. Select an input for the rule statement.
   - If the rule specification does not contain an input that you can use, create an input. Configure the input properties to represent the type of column that contains the business data.
5. Select an operator to validate the results of the condition analysis.
   - To specify that the input data must match the data that you enter, select the following operator:

     `contains`
6. Select the type of condition to apply to the input.
   - To compare the input data to a string value, select the following condition type:

     `string value`
7. Enter the value to compare to the input data. For example, enter "West."

   The condition searches the rule statement input for the string value that you enter.

Note: You can also compare the rule statement input to a value in another input column on the same row. When you specify an input column as the condition type, the condition looks for the condition input data in the rule statement input data.

8. Select the type of action to apply to the data that satisfies the condition logic. For example, you might configure an action that returns the inventory numbers for each product.
   - To return the data values from another column on the same row, select the following action type:

     ```
     input
     ```
9. Select the input that represents the inventory level for each product in the warehouse.
10. Save the rule specification.

CHAPTER 6

Test and Validation Operations

This chapter includes the following topics:

- Test and Validation Operations Overview, 55
- Rules and Guidelines for Validation Operations, 55
- Validating a Rule Specification, 56
- Testing a Rule Specification, 56
- Testing a Rule Set, 57

Test and Validation Operations Overview

Validate a rule specification to verify that the rule specification is valid for use in the Developer tool. For example, validate a rule specification before you generate mapplets from the rule specification.

Test the rule specification to verify that the output satisfies the business rule requirements. To test a rule specification or a rule set, you enter sample values in one or more data columns. Use the Test options on the rule specification properties to display the columns. You can cut, copy, and paste test data within the test area and between rule sets in the rule specification. If you open more than one rule specification in concurrent browser tabs, you can cut, copy, and paste test data between the rule specifications.

You can test a complete rule specification, and you can test each rule set in the rule specification. If a rule specification is not valid, test the rule sets to find and fix the validation error.

Rules and Guidelines for Validation Operations

Before you release a rule specification to the Developer tool users, or before you generate a mapplet from a rule specification, validate the rule specification.

Consider the following rules and guidelines when you validate a rule specification:

- The Analyst tool validates the lowest rule sets in the rule specification first. If the Analyst tool finds an error in a rule set, the Analyst tool stops validation and reports the error.
- The Analyst tool reports one error at a time. When you repair a validation error, validate the rule specification again.

- A rule specification can fail validation for the following reasons:
  - A condition or an action in a rule statement specifies a data type that differs from the input data type.
  - The rule statements in a rule set generate output data values of different data types.
  - A parent rule set does not read an output from a child rule set.
  - A rule statement is incomplete.
- When you move or copy a rule set to another rule specification, you move or copy any test data that you added for the rule set. You also move or copy any input that the rule set uses.

Validating a Rule Specification

Validate a rule specification to verify that the rule specification is ready for use in the Developer tool. Additionally, validate a rule specification to verify that you can generate the mapplets that the rule specification defines.

1. Open the rule specification.
2. Click **Validate**.

   The rule specification highlights any rule set that is not valid. Click a rule set to read a message that describes the validation error.

Testing a Rule Specification

Test a rule specification to verify that the data flows through rule specification in the ways that you expect. If possible, use the business data to test the rule specification.

1. Open the rule specification.
2. Select the top-level shape in the rule specification.
3. In the rule specification properties, click **Test**.

   The properties view displays the input columns that you configured in the rule specification.
4. Enter one or more data values in an input column. You can enter data values for any rule set in the rule specification.

   If you enter date data in an input column, use one of the following date formats:

   - yyyy-MM-dd
   - yyyy-MM-dd hh:mm:ss
   - yyyy-MM-dd hh:mm:ss.SSS

   Note: To test a rule statement that reads a list of date values as an input, enter the date data in the format yyyy-MM-dd.
5. Click **Test Rule**.

   The properties view displays the path of the data values through the rule specification.
6. Verify that the rule sets read the data values and write the output values that you expect.

- A data value from a child rule set might generate output values on multiple rule sets. Verify that the test results indicate a data value on each rule set.
- If a data value does not generate an output on a rule set, the test results show an empty field for the rule set.

Testing a Rule Set

Test a rule set to verify that the rule statements generate the correct output values. If possible, use the business data to test the rule specification.

To verify the full operation of the rule specification, select the top-level rule shape and test the rule specification.

1. Open the rule specification.
2. Select a rule set.
3. In the rule set properties, click **Test**.

   The properties view displays the input columns that you configured in the rule set.
4. Enter one or more data values in an input column. You can enter data values for any rule set in the rule specification.

   If you enter date data in an input column, use one of the following date formats:

   - yyyy-MM-dd
   - yyyy-MM-dd hh:mm:ss
   - yyyy-MM-dd hh:mm:ss.SSS

   Note: To test a rule statement that reads a list of date values as an input, enter the date data in the format yyyy-MM-dd.
5. Click **Test Rule**.

   The properties view displays the path of the data values through the rule specification.

INDEX

A

Analyst Service
  mapplets 11

B

business rules
  facts 21

C

conditions
  Define Action dialog box 33
  Define Condition dialog box 29
copy and paste operations
  rule sets 25
  rule statements 40

D

Data Integration Service
  current date analysis 28
  mapping execution 11
  mappings and mapplets 11
date elements in functions 36
Developer tool 11

F

facts in business data 21
function expressions
  Add to Date 34
  Choose 34
  Concatenate 34
  Date Difference 34
  Date Part 34
  Date Time 34
  Greatest 34
  Last Day 34
  Least 34
  Length 34
  Lower 34
  Null 34
  Replace Character 34
  Replace String 34
  Reverse 34
  Truncate 34
  Upper 34
functions
  defining a range of values 50

G

generating rules
  rules and guidelines 19
  steps to generate rules 19

I

inputs
  data types 10

M

Manage Global Inputs dialog box 18
mappings
  Data Integration Service 11
mapplets
  adding to a rule statement 50
  Analyst Service 11
  Data Integration Service 11
  push down to hadoop 11
  synonym for rules 11
Model repository 11

N

naming conventions 11

R

rule sets
  adding a rule set to a rule specification 25
  configuration 21
  configuration rules and guidelines 21
  copy and paste operations 25
  inputs 10
  properties 24
  testing a rule set 57
rule specifications
  collaborating with developers 11
  configuration prerequisites 14
  configuring a rule specification 17
  general properties 16
  generating a mapplet from a rule specification 19
  inputs 10
  mapplets and mappings 11
  Model repository 11
  push down to hadoop 11
  rule sets 10, 21
  rule statements 10, 27
  rules and guidelines for mapplets and rules 11
  terminology 11

rule specifications *(continued)*
testing a rule specification 56
verify business data properties 15
verify business rule requirements 15
version control 12
rule statements
actions 10, 31
add rule statement using result 39, 50
AND relationships 47
conditions 10, 28
configuration 27
copy and paste operations 40
current date analysis 28
default rule statement 10
defining a range of values 50
discovering business information 47
function expression descriptions 34
improving data usability 46
List of Values function 23
menu options 37
null value analysis 28
operators 10, 30
rule statements *(continued)*
searching a list of values 52
standardizing data 45
steps to configure 38
verify data against business regulations 48
verifying data accuracy 44
rules
synonym for mapplets 11

T

transformations 11

V

version control
generating mapplets 19
rule specifications 12