Cleo Integration Cloud | *Clarify*

Chapter | # Transforming multi-format flat files

Overview | This topic demonstrates how to transform a delimited flat file to a fixed-length flat file.

Both are multi-format.

Features, functions, and tips not discussed in previous demos include: multiple Composite Rules, Rule Conditions, user-defined variables, and copying rules within a Ruleset.

Contents |

Scenario

The source data is generated by a back-end application that aggregates incoming purchase orders as a pipe-delimited multi-format flat file.

The actual source data to be transformed is for three purchase orders, as illustrated here.

```
PO|BestValue_22|1/1/2013
CI|HI-BVGS|BEST VALUE GOODSTUFF|1 BEST VALUE BLVD||HONOLULU|HAWAII|96815
ST|Same as Company
BT|Same as Company
CN|BOB SMITH|(555) 555-5555|BSMITH@BVGS.COM
SI|
LI|1|PINGBALL|WHITE PING PONG BALLS (100)|TABLE GAME ITEMS|500682|200|Case|16|3200
LI|2|BARSOAP|DECORATIVE SOAP BAR|KEEP IT CLEAN|500227|900|Each|0.75|675
TO|2|1100|3875
PO|GoodStuff 303|1/2/2013
CI|HI-BVGS|BEST VALUE GOOD STUFF|1 BEST VALUE BOULEVARD||HONOLULU|HAWAII|96815
ST|New|ABC DISTRIBUTION|1234 MAIN STREET|PO BOX 938|CLEVELAND|OHIO|44109-1234
BT|Same as Shipping
CN|SUSAN JOHNSON|(555) 555-6666|SJOHNSON@BVGS.COM
SI|Please note the new SHIPTO address
LI|1|LMWG01|LARGE MANS WORK GLOVES BLACK|GLOVELY STUFF|500151|160|Pair|9.25|1480
LI|2|LMWC01|LARGE MANS WINTER COAT BLACK|MANITZ COLDOUT|500155|6|Each|100|600
LI|3|SWG-RED|SMALL WOMENS GLOVES - RED|GLOVELY STUFF|500001|24|Pair|12.5|300
TO|3|190|2380
PO|FV-2013_0047|1/3/2013
CI|WY-FV|FAMILIAR VASES|100 INDUSTRIAL WAY|SUITE 123|CASPER|WYOMING|82640
ST|Same as Company
BT|New|BOB SMITH|23 SOUTH CENTER ST||POTTSVILLE|PENNSYLVANIA|17901
CN|BOB SMITH|(555) 555-5555|BSMITH@BVGS.COM
SI|New Billing Address|Please handle with care|
LI|1|PV-28 EQ|28-INCH PORCELAIN VASE|FANCY VASES||1|Each|375|375
TO|1|1|375
```

Each of its eight formats has its own set of fields.

Format	Data Type	Fields
PO	Header	POnumber, POdate
CI	Company	CIcode, CIname, CIadd1, CIadd2, CIcity, CIstate, CIzip
ST	Ship To	STcode, STname, STadd1, STadd2, STcity, STstate, STzip
BT	Bill To	BTcode, BTname, BTadd1, BTadd2, BTcity, BTstate, BTzip
CN	Contact	CN, CNname, CNphone, CNemail
SI	Instructions	SInote1, SInote2, SInote3
LI	Line Items	LInumber, LIyourSKU, LIdesc, LIbrand, LIourSKU, LIqty, LIunit, LIcost, LIextcost
TO	Totals	TOlines, TOquantity, TOcost

The target fixed-length multi-format flat file provides two types of information.

First, it indicates any new billing/shipping addresses in the source data. It then presents each purchase order's total number of line items, quantity ordered, and price.

As only two addresses are indicated as new and there are three purchase orders in the source data, the expected target data is illustrated here.

```
NUM00000000000000000002000000000000000001100
00000000000000003875
TXTHI-BVGS          ST              ABC DISTRIBUTION   1234 MAIN STREET    PO BOX 938
CLEVELAND           OHIO            44109-1234
NUM00000000000000000003000000000000000000190
00000000000000002380
TXTWY-FV            BT              BOB SMITH          23 SOUTH CENTER ST
POTTSVILLE          PENNSYLVANIA    17901
NUM00000000000000000001000000000000000000001
00000000000000000375
```

Resources

The resources needed to accomplish this transformation are meant to provide a basic understanding of how Clarify objects work together. All resources are created and defined on the *Clarify Workbench | Project Explorer*.

- Project [**com.training.demo02.multifftoff**] – A Project is a directory within the Workspace.

- Package [**com.training.demo02.multifftoff**] – A Package is a directory within a Project.

- File Adapter [**ReadTxtFA**] – This File Adapter brings the source data to Clarify.

- File Adapter [**WriteTxtFA**] – This File Adapter sends the target data out of Clarify.

- Flat File Schema [**DelimitedFF**] – This Schema describes the delimited flat file data.

- Flat File Schema [**FixedFF**] – This Schema describes the fixed-length flat file data.

- Transformation Settings [**PipeDelimitedSourceTS**] – This Transformation Settings object informs the Business Process of the delimiter to use when reading the source file. No delimiter is necessary for writing the target file, as it has a fixed-length format.

- Ruleset [**DelimitedFfToFixedFfRS**] – This Ruleset transforms the data from the source format to the target format.

- Business Process [**DelimitedFfToFixedFfBPS**] – This Business Process sequences three user-defined objects to read the source data, transform it to the target format, and write that out to a file on the local file system.

Project: com.training.demo02.multifftoff

Select *File | New | EBI Project*.

Type **com.training.demo02.multifftoff** in the *Project name* field.

Click **Finish**.

When Clarify creates a Project, it automatically creates a Package named the same as the Project. All resources for this Project are created in its *com.training.demo02.multifftoff* Package.

File Adapter: ReadTxtFA

Select the Package.

Select *File | New | File Adapter*.

Choose *Read* and click **Next**.

Type **ReadTxtFA** in the *Name* field.

Click **Finish**. The object is created, and its editor appears.

Click in the *File Name* field. Click the **Open** button, locate/select the *source_data.txt* file in the *C:\Clarify_demos\demo02\in* folder, and click **Open**.

Save the object and close the editor.

File Adapter: WriteTxtFA

Select the Package.

Select *File | New | File Adapter*.

Choose *Write* and click **Next**.

Type **WriteTxtFA** in the *Name* field.

Click **Finish**. The object is created, and its editor appears.

Click in the *File Path* field. Click the **Browse For Folder** button, locate/select the *C:\Clarify_demos\demo02\out* folder, and click **OK**.

Type **target_data.txt** in the *File name* field.

Save the object and close the editor.

Schema: DelimitedFF

To create a Schema from a sample file, the sample must reside in the Workspace.

Bring the sample file to Clarify

Right-click the Project and select *New | Folder*.

Type **samples** in the *Folder name* field and click **Finish**.

Copy the *sample_source_data.txt* file in the *C:\Clarify_demos\demo02* folder.

> This sample file has one instance of each record format, with data in each of the fields for every format, as illustrated here.
>
> ```
> PO|PONumber_123|1/1/2017
> CI|HI-BVGS|BEST VALUE GOOD STUFF|1 BEST VALUE BOULEVARD||HONOLULU|HAWAII|12345-6789
> ST|New|ABC DISTRIBUTION COMPANY|1234 MAIN STREET|PO BOX 938|CLEVELAND|OHIO|44109-1234
> BT|New|CHUCK BUCHANAN|23 SOUTH CENTER ST|ADDRESS LINE 2|POTTSVILLE|PENNSYLVANIA|17901-9876
> CN|CHUCK BUCHANAN|(555) 555-5555|BSMITH@BVGS.COM
> SI|Note New Billing and Shipping Addresses|Please handle with care|
> LI|1|BARSOAP|DECORATIVE SOAP BAR|KEEP IT CLEAN|500227|900|Each|0.75|675.23
> TO|1|900|675
> ```

Right-click the Project's *samples* folder and select *Paste*.

Create the Schema
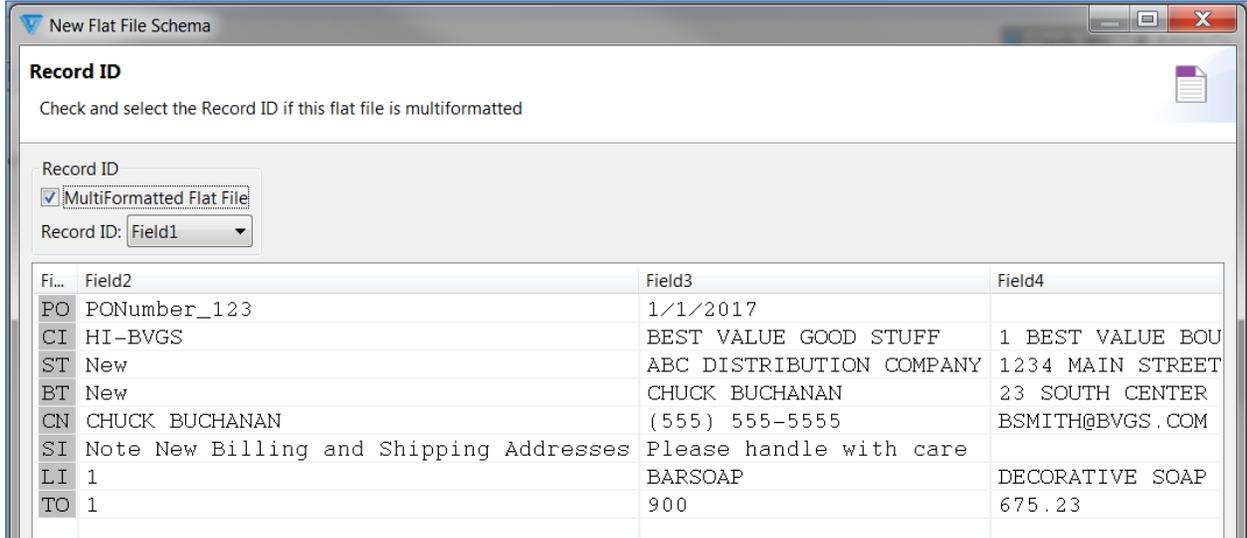
Select the Package.

Select *File | New | Flat File Schema*. A wizard assists with creating and defining this object.

On the *Create* panel, choose *Create From Sample* and click **Next**.

On the *Create From Sample* panel, click **Browse**, select the *sample_source_data.txt* file, and click **OK**. Click **Next**.

On the *Fixed Length or Delimited* panel, as Clarify has correctly determined that this file is pipe-delimited, click **Next**.

On the *Record ID* panel, select the *MultiFormatted Flat File* checkbox.

Clarify correctly assumes that *Field1* is the Record ID, so click **Next**.

The *Record Structure* panel is where the fields of each record may be renamed.

One at a time, select each record node in the *Record Structure* area, and in the *Field Properties* area provide a new *Field Name* for each field, as indicated below. Click the **Next Field** button to advance through the fields.

Default	Field Name
PO	POformat, POnumber, POdate
CI	CIformat, CIcode, CIname, CIadd1, CIadd2, CIcity, CIstate, CIzip
ST	STformat, STcode, STname, STadd1, STadd2, STcity, STstate, STzip
BT	BTformat, BTcode, BTname, BTadd1, BTadd2, BTcity, BTstate, BTzip
CN	CNformat, CNname, CNphone, CNemail
SI	SIformat, SInote1, SInote2, SInote3
LI	LIformat, LInumber, LIyourSKU, LIdesc, LIbrand, LIourSKU, LIqty, LIunit, LIcost, LIextcost
TO	TOformat, TOlines, TOquantity, TOcost

Click **Next**.

The *Transformation Settings* panel allows for automatic creation at this time of the Transformation Settings object. Choose *None* (as we'll create it manually later) and click **Next**.

The *Flat File Schema* panel allows for specifying the object's location and name. Type **DelimitedFF** in the *Name* field and click **Finish**. The object is created and its editor appears.

Click the *RecordGroup0* node. In the *Properties* view's *Properties* tab, click in the *Name* field, type **Source**, and click **OK**.
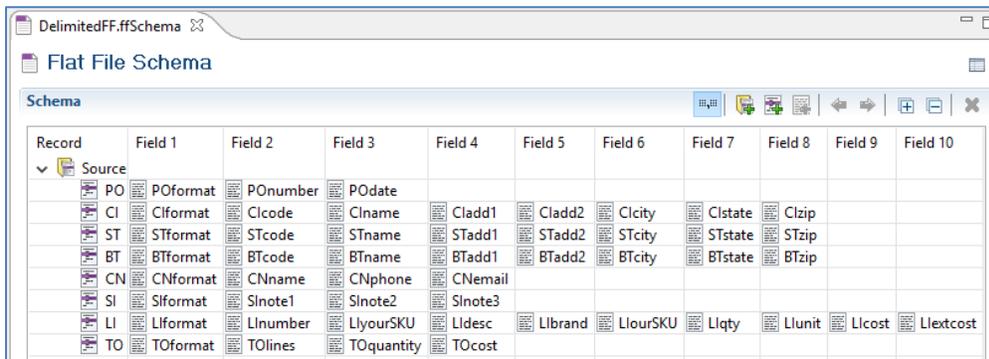
Save the object and close the editor.

Schema: FixedFF

To create a Schema from a sample file, the sample must reside in the Workspace.

Bring the sample file to Clarify

Copy the *sample_flatfile_with_20_fields_fixed.txt* file in the *C:\Clarify_demos\demo02* folder.

Right-click the Project's *samples* folder and select *Paste*.

Create the Schema

Select the Package.

Select *File | New | Flat File Schema*. A wizard assists with creating and defining this object.

The *Create* panel allows for using a sample flat file or creating an empty Schema. Choose *Create From Sample* and click **Next**.

The *Create From Sample* panel allows for selecting the flat file to use. Click **Browse**, select the *sample_flatfile_with_20_fields_fixed.txt* file, and click **OK**. Click **Next**.

The *Fixed Length or Delimited* panel allows for changing the detected *Flat File Type* and *Delimited Settings*. As Clarify has correctly determined that this is a fixed-length file, no changes are necessary. Click **Next**.

The *Alter Fields* panel allows for inserting, removing, and resizing fields in a fixed-length flat file. As the only alteration is to separate the values in the *Field1* column, which is already selected, click the **Add** button to insert a column to its right. This new column is named *Field11*. With the *Field1* column still selected, change its *Field Length* value to **3**, which is the length of the format names (*TXT* and *NUM*). All of the characters originally in that field from the fourth position to the end now adjust themselves properly. Click **Next**.

The *Record ID* panel allows for indicating this Schema as multi-format and, if so, which column serves as the record ID. As this is a multi-format Schema, select the *MultiFormatted Flat File* checkbox. Clarify correctly assumes that *Field1* is the Record ID, so click **Next**.

The *Record Structure* panel is where the fields of each record may be renamed. No renaming is necessary for this generic Schema, so click **Next**.

The *Flat File Schema* panel allows for specifying the object's location and name. Type **FixedFF** in the *Name* field and click **Finish**. The object is created, and its editor appears.

Click the *RecordGroup0* node. In the *Properties* view's *Properties* tab, click in the *Name* field, type **Target**, and click **OK**.

Save the object and close the editor.

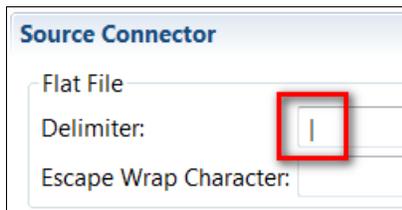## Transformation Settings: PipeDelimitedSourceTS

Select the Package.

Select *File | New | Transformation Settings*.

Type **PipeDelimitedSourceTS** in the *Name* field.

Click **Finish**. The object is created, and its editor appears.

Type a pipe in the *Delimiter* field of the *Source Connector* section.

Save the object and close the editor.

Ruleset: DelimitedFfToFixedFfRS

This Ruleset transforms the data from the source format to the target format.

Select the Package.

Select *File | New | Ruleset*.

Type **DelimitedFfToFixedFfRS** in the *Name* field.

Click **Next**. A wizard assists with defining this object.

The *Schema Types* panel allows for selecting both the source and target Schemas. Choose *Flat File* for each. Click **Next**.

The *Source Schema* panel allows for selecting an existing Schema. Click **Browse**, select the *DelimitedFF* object, and click **OK**. Click **Next**.
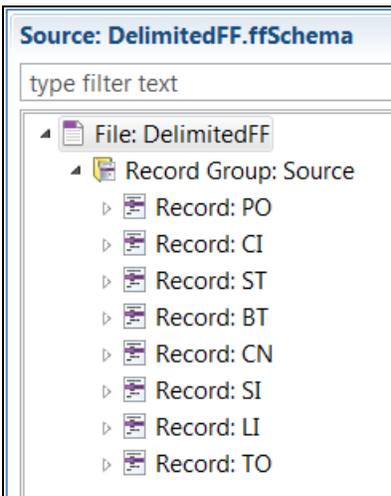
The *Target Schema* panel allows for selecting an existing Schema. Click **Browse**, select the *FixedFF* object*,* and click **OK**.
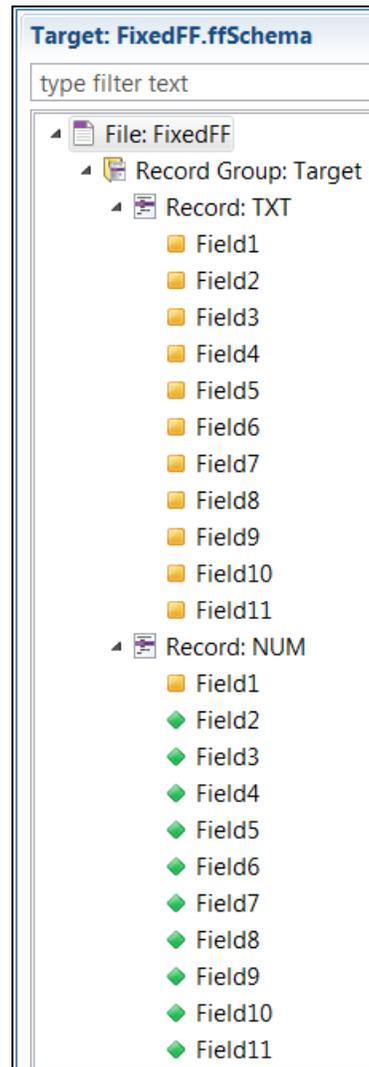
Accept **Version 2** for the *Ruleset Version*.

Click **Finish**.

The object is created, and its editor appears.

Expand the *Source* section's *File: DelimitedFF* node and then its *Record Group: Source* node to display all record nodes.

Click the **Expand All** button in the *Target* section to display all levels.

Source: DelimitedFF.ffSchema

type filter text

- ▲ 📄 File: DelimitedFF
  - ▲ 📑 Record Group: Source
    - ▷ 📰 Record: PO
    - ▷ 📰 Record: CI
    - ▷ 📰 Record: ST
    - ▷ 📰 Record: BT
    - ▷ 📰 Record: CN
    - ▷ 📰 Record: SI
    - ▷ 📰 Record: LI
    - ▷ 📰 Record: TO

Target: FixedFF.ffSchema

type filter text

- ▲ 📄 File: FixedFF
  - ▲ 📑 Record Group: Target
    - ▲ 📰 Record: TXT
      - 🟨 Field1
      - 🟨 Field2
      - 🟨 Field3
      - 🟨 Field4
      - 🟨 Field5
      - 🟨 Field6
      - 🟨 Field7
      - 🟨 Field8
      - 🟨 Field9
      - 🟨 Field10
      - 🟨 Field11
    - ▲ 📰 Record: NUM
      - 🟨 Field1
      - 🔶 Field2
      - 🔶 Field3
      - 🔶 Field4
      - 🔶 Field5
      - 🔶 Field6
      - 🔶 Field7
      - 🔶 Field8
      - 🔶 Field9
      - 🔶 Field10
      - 🔶 Field11

This Ruleset requires three Composite Rules and several child Simple Rules for each. The first two Composite Rules must be conditioned.

Composite Rule 1: Record ST

To create and define this Composite Rule, drag the source's *Record: ST* node and drop it on the target's *Record: TXT* node. This rule, which relates the source and target Schemas at the respective record levels, appears in the *Rules* section.

Condition Composite Rule 1

This Composite Rule must be conditioned to execute only when the source data indicates that the Ship To address is new.

Select the Composite Rule and display the *Properties* view's *Condition* tab.

Display the Ruleset editor's *Actions* panel, type **str** in the *Filter* field to restrict the display to actions containing that text and drag *StringEquals* to the *Condition* tab's *Condition* field.

Drag the source's *ST* record's *STcode* node to the *Condition* tab's *SourceString1* property.

Type **New** in the *Condition* tab's *SourceString2* property and press **Enter**.

Simple Rules for Composite Rule 1

Ensure that this Composite Rule is selected before creating its child Simple Rules, as indicated here.

Source Record: Field	Target Record: Field
Value: TXT	TXT: Field1
CI: CIcode	TXT: Field2
ST: STformat	TXT: Field3
ST: STname	TXT: Field4
ST: STadd1	TXT: Field5
ST: STadd2	TXT: Field6
ST: STcity	TXT: Field7
ST: STstate	TXT: Field8
ST: STzip	TXT: Field9

To create the first rule, click the **New Rule** button. In the list of available actions, type **mov** in the *Filter* field to restrict the display to actions containing that text. Double-click *Move* to select that action. The undefined rule appears in the *Rules* section.

To define the first Rule, display the Properties view's *Rule* tab, type **TXT** in the *sourceNode* property and press **Enter**. Re-select this partially-defined rule and drag the target's *TXT: Field1* node to the *Return Assignments* area.

To create and define the remaining rules, drag the source Schema node to the target Schema node.

Composite Rule 2: Record BT

To create this Composite Rule, right-click the previous Composite Rule and select *Add sibling | Composite Rule | Composite Rule*. This currently undefined rule, which is to relate the source and target Schemas at the *BT* and *TXT* record levels, respectively, appears in the *Rules* section.

To define this Composite Rule, drag the source's *Record: BT* node to the *Properties* view's *Source* property, and then drag the target's *Record: TXT* node to the *Target* property.

Condition Composite Rule 2

This Composite Rule must be conditioned to execute only when the source data indicates that the Bill To address is new. For this Composite Rule, a user-defined variable will be used rather than the literal "New" as before.
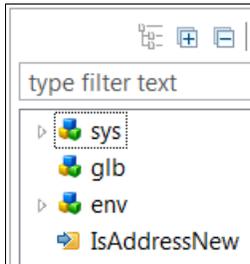
Display the Ruleset editor's *Variables* panel. Click the **Add** button.
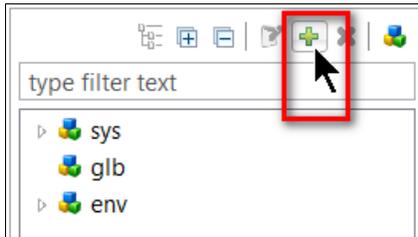
In the Create a New Variable window, type **IsAddressNew** in the *Name* field, select *String* as the *Type*, and type **New** in the *Value* field.

Click **OK**.

The variable appears at the bottom of the list.

Select the Composite Rule and display the *Properties* view's *Condition* tab.

Display the Ruleset editor's *Actions* panel and drag *StringEquals* to the *Condition* tab's *Condition* field.

Drag the source's *BT* record's *BTcode* node to the *Condition* tab's *SourceString1* property.

Display the Ruleset editor's *Variables* panel and drag *IsAddressNew* to the *Condition* tab's *SourceString2* property.

Simple Rules for Composite Rule 2

Ensure that this Composite Rule is selected before creating its child Simple Rules, as indicated here.

Source Record: Field	Target Record: Field
Value: TXT	TXT: Field1
CI: CIcode	TXT: Field2
BT: BTformat	TXT: Field3
BT: BTname	TXT: Field4
BT: BTadd1	TXT: Field5
BT: BTadd2	TXT: Field6
BT: BTcity	TXT: Field7
BT: BTstate	TXT: Field8
BT: BTzip	TXT: Field9

As the first two rules above are identical to the first two rules of the previous Composite Rule, they may be copied.

To create these first two rules, multi-select the first two rules of the <u>first</u> Composite Rule, right-click, select *Copy*.
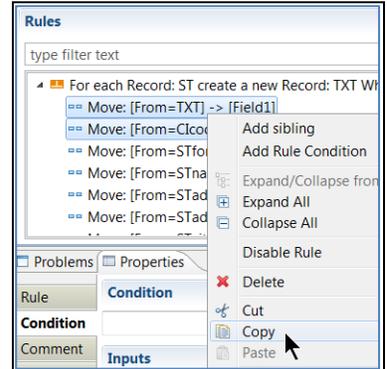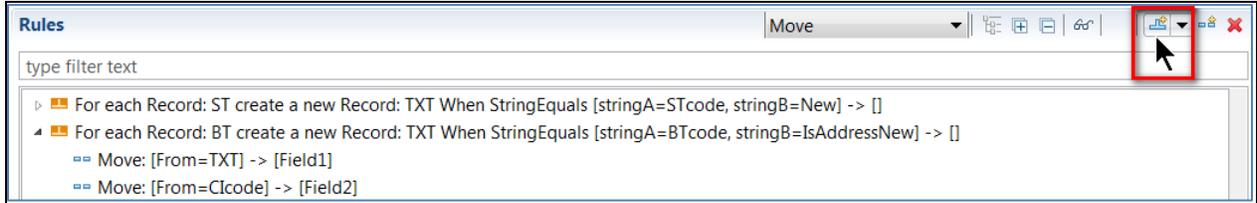
Right-click <u>this</u> Composite Rule and select *Paste*.
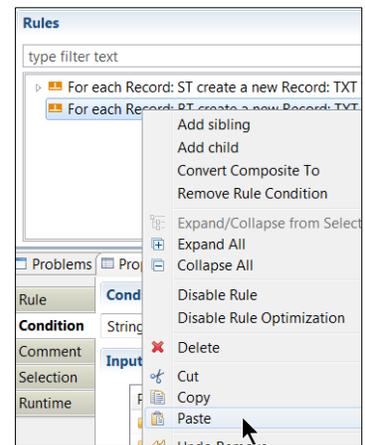
To create and define the remaining rules, drag the source Schema node to the target Schema node.

Composite Rule 3: Record TO

To create this Composite Rule, right-click the first (top-most) Composite Rule and click the **New Composite Rule** button.

This currently undefined rule, which is to relate the source and target Schemas at the *TO* and *NUM* record levels, respectively, appears in the *Rules* section.
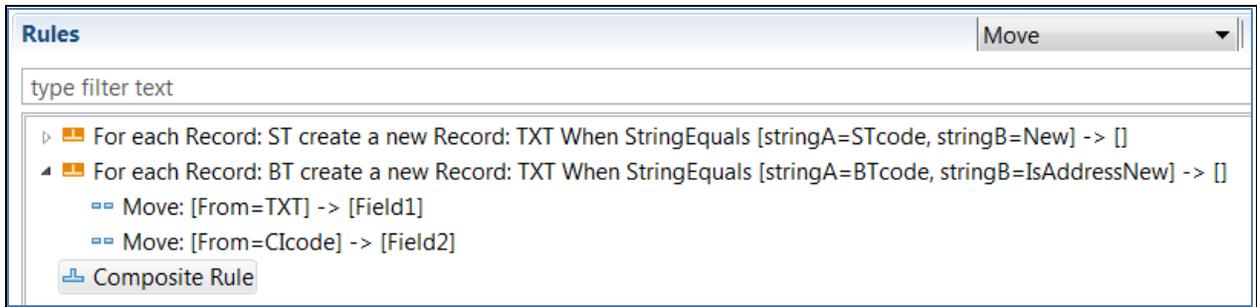
To define this Composite Rule, display the *Properties* view's *Rule* tab, then drag the source's *Record: TO* node to the *Properties* view's *Source* property, and then drag the target's *Record: NUM* node to the *Target* property.

Simple Rules for Composite Rule 3

Ensure that this Composite Rule is selected before creating its child Simple Rules, as indicated here.

Source Record: Field	Target Record: Field
Value: NUM	NUM: Field1
TO: TOlines	NUM: Field2
TO: TOquantity	NUM: Field3
TO: TOcost	NUM: Field7

To create the first rule, click the **New Rule** button. In the list of available actions, type **mov** in the *Filter* field to restrict the display to actions containing that text. Double-click *Move* to select that action. The undefined rule appears in the *Rules* section.

To define the first rule, display the *Properties* view's *Rule* tab, drag the target's *NUM: Field1* node to the *Return Assignments* area, and type **NUM** in the *From* property and press **Enter**.

To create and define the next three rules, drag the source Schema node to the target Schema node.

Save the object and close the editor.

Business Process: DelimitedFfToFixedFfBPS

Select the Package.

Select *File | New | Business Process*.

Choose *No Template* and click **Next**.

Type **DelimitedFfToFixedFfBPS** in the *Name* field.

Click **Finish**. The object is created, and its editor appears.

This Business Process requires three tasks to (1) read the source data, (2) transform the data, and (3) write the target data.

Task 1: Read the source data

To create this task, click the **Add** button in the *Script* section. Click that step's *Click to select task* entry. In the list of available tasks, type **read** in the *Filter* field to restrict the display to tasks containing that text. Double-click the *ReadTxtFA* user-defined object to select it. That task is displayed on the top line.

To define this task, click the *Properties* view tab.

This task has one parameter.

Click in the *storageNodes* parameter's field. This parameter represents the AdapterPayload (the data of the source file). The bold-italic text indicates that this is a mandatory parameter.

A list of available parameters and variables appears. As none yet exist, that list is empty.

Click **New Variable**. An AdapterPayload may be an Array, Object, or StorageNode. Double-click *StorageNode* to indicate it as the type for this variable.

newVariable appears in the *Properties* view and also appears in the *Variables* section, with its name pre-selected, ready for a name change.

Type **source** and press **Enter** to rename this variable. This new name is reflected in the *Properties* view.

Task 2: Transform the data

To create this task, click the **Add** button in the *Script* section. Click that step's *Click to select task* entry. In the list of available tasks, type **rs** in the *Filter* field to restrict the display to tasks containing that text. Double-click the *DelimitedFfToFixedFfRS* user-defined object to select it. That task is displayed on the second line.

To define this task, click the *Properties* view tab.

This task has several parameters. Provide values for the first three; the last three are not necessary here.

Click in the *sourceNode* parameter's field. This parameter represents the StorageNode containing the data to be transformed. In the list of available parameters and variables, double-click the previously-defined *source* variable, as the data read by the first task is the data to be transformed.

Click in the *targetNode* parameter's field. This parameter represents the StorageNode containing the transformed data. A list of available parameters and variables appears. Click **New Variable**. *newVariable* appears in the *Properties* view and also appears in the *Variables* section, with its name pre-selected, ready for a name change. Type **target** and press **Enter** to rename this variable. This new name is reflected in the *Properties* view.

Click in the *settings* parameter's field. This parameter represents the Transformation Settings object to be used during transformation. Click the *TransformationSettings Literal* tab and double-click the previously-defined *PipeDelimitedSourceTS* object to select it.

Task 3: Write the target data

To create this task, click the **Add** button in the *Script* section. Click that step's *Click to select task* entry. In the list of available tasks, type **wri** in the *Filter* field to restrict the display to tasks containing that text. Double-click the *WriteTxtFA* object to select it. That task is displayed on the third line.

To define this task, click the *Properties* view tab.

This task has only one parameter.

Click in the *storageNodes* parameter's field. This parameter represents the AdapterPayload (the data to be written to the target file). In the list of available parameters and variables, double-click the previously-defined *target* variable, as the data transformed by the second task is the data to be written.

Save the object and close the editor.

Deploy, Launch, Audit

Now that all of the Project's resources are created on the *Clarify Workbench | Project Explorer*, the remaining activities take place on various views of the *Admin Console* perspective.

The Project's top-level object – in this case, the Business Process – must be deployed to the *Local Test* server. Then, that Business Process can be launched and the result audited.

Deploy

Access *Admin Console | Projects*.

In the *Workspace* section, expand the *com.training.demo02.multifftoff* Project, and then expand its Package node.

Drag the *DelimitedFfToFixedFfBPS* object to the *Selected Scenarios* section.

Start the *Local Test* server by ensuring that it is displayed in the *Server Environment* dropdown and then click the **Start** button.

When fully-expanded, the *Server Projects* section shows all installed objects.
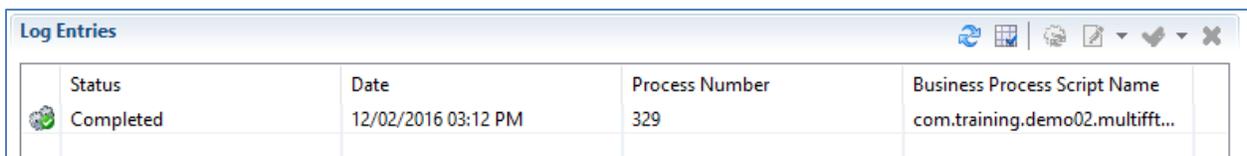
Launch

Access *Admin Console | Business Process*.

Select *com.training.demo02.multifftoff.DelimitedFfToFixedFfBPS* and click the **Launch** button.

Audit

Access *Admin Console | Auditor*.

The *Log Entries* section shows the results.

	Status	Date	Process Number	Business Process Script Name
🐸	Completed	12/02/2016 03:12 PM	329	com.training.demo02.multifft...