Grouper Element

The element groups data supplied to the specified slot by the specified property (for example, by value). Additionally, it is possible to merge data from another slots associated with the specified one.

Element type: grouper

Parameters

To use the *Grouper* element connect the *Grouper's* input port to the required workflow element. Select the *Grouper* element on the *Scene* and specify *Group slot* and *Group operation* parameters in the *Parameters* area in the *Property Editor*. To merge associated data, it is possible to create as many *Output slot (s)* as required (see details below).

Group slot

The *Group slot* specifies a *slot* that is used to group the input data. The list of available values of the parameter depend on the slots of workflow elements which produce data in the workflow before the *Grouper* element. There is a special *Unset* value. When it is selected, only one group is created.

Group operation

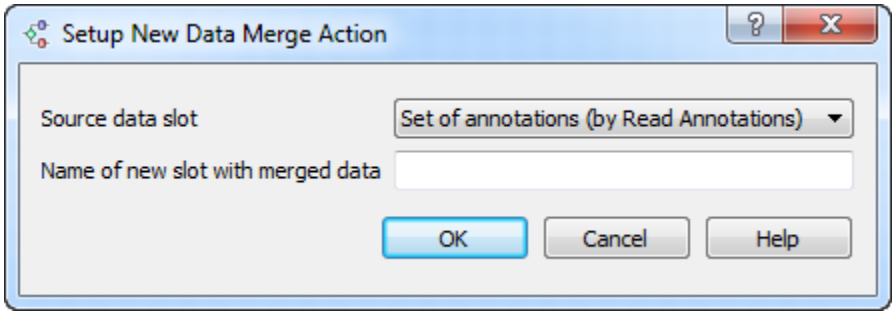
The *Group operation* specifies criteria to group data supplied to the *Group slot*. It can take the following values:

- *By value* - input data are compared by value (a group is created for each unique value, it can contain one or several identical values)
- *By identity* - input data are compared by internal data ID (all values are unique)
- *By name* - input data are compared by their names

By value group operation is available for group slots of types *Sequence*, *Set of annotations*, *MSA*, *Plain text*, *Source URL*. *By identity* and *By name* group operations are available for group slots of type *Sequence* only.

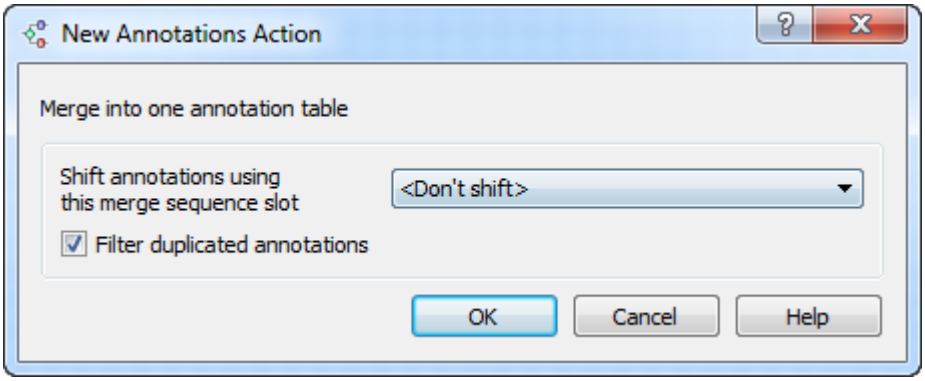
Output slots

When data supplied to the *Group slot* are divided into different groups the associated data are also got into a group. The possible associated data depend on the workflow. For example, a *Sequence Reader* element contains slots *Sequence* and *Set of annotations*. These data are **associated** as annotations belong to a sequence. Another example of associated data are sequence markers created by the *Sequence Marker* element. The associated data, therefore, can be additionally handled (i.e. merged) by the *Grouper*element. The action that can be performed on the associated data depends on their type. In any case to output handled associated data you need to create a new output slot in the*Grouper* element. To create it click the *Add* button in the *Grouper's Parameters* area. The following dialog appears:



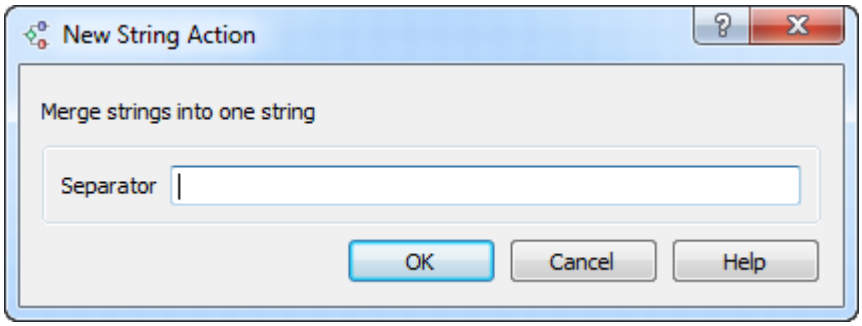
In the dialog you should select a *Source data slot* (i.e. a slot with the associated data) and input a name of the new slot. Click the OK button. A new dialog appears that specifies how the associated data should be merged. The view of the dialog and the available merge actions for different types of the *Source data slot* are the following:

- For a *Set of annotations* slot the *New Annotations Action* dialog appears:



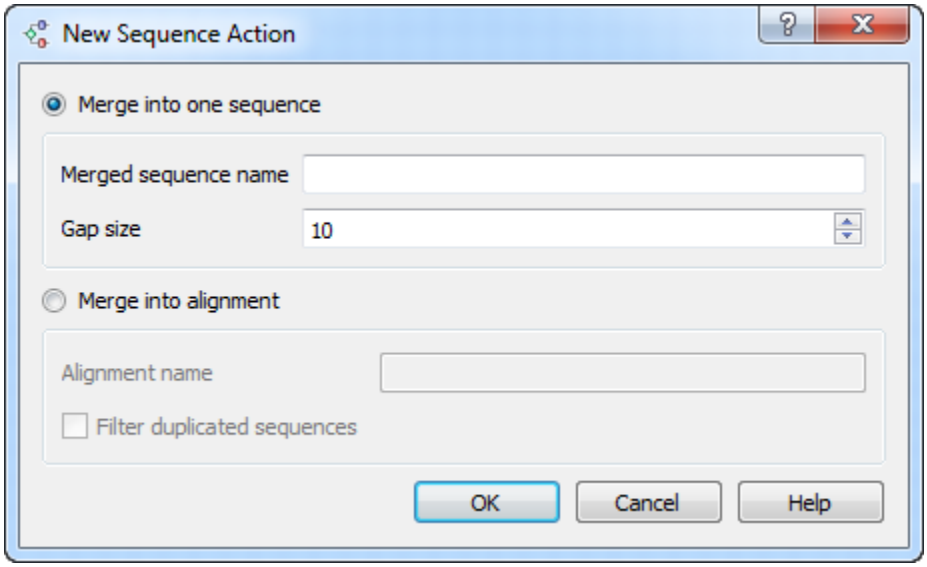
You can merge annotations into one annotation table and, optionally, filter duplicated annotations. Also, you can shift annotations. To do it, you need to create another output slot with type *Sequence* and *Merge into one sequence* option selected (see below). In other words you need to merge all sequences in a group into one sequence. In this case you select the corresponding sequence slot in the *New Annotations Action* dialog and each set of annotations in a group is shifted according to the corresponding sequence in the group. As the result you have one sequence and one set of annotations allocated on the whole sequence.

- For *Source URL* and *Plain text* slots the *New String Action* dialog appears:



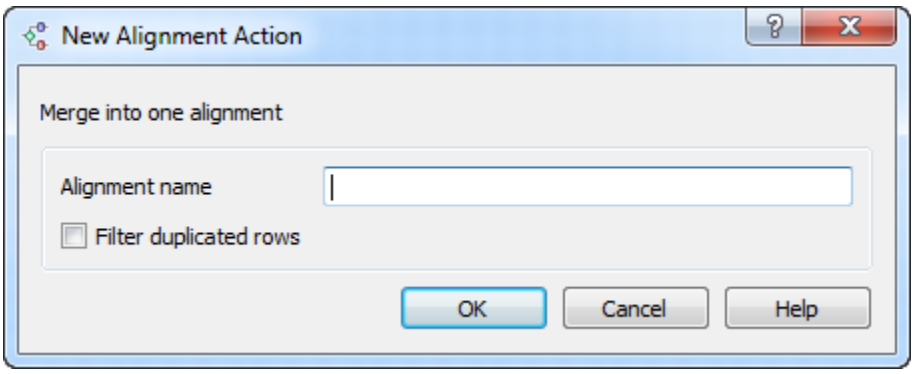
Using this dialog you can merge strings into one string. Optionally, you can specify an additional strings separator.

- For a *Sequence* slot the *New Sequence Action* dialog appears:



You can either merge all sequences in a group into one sequence or create a multiple sequence alignment. In the first case you need to specify the *Merged sequence name* and you can select the number of unknown characters between the merged sequences. In the second case you need to specify the alignment name. To filter duplicated sequence check the corresponding check box.

- For a *MSA* slot the *New Alignment Action* dialog appears:



Input the alignment name in this dialog. To filter duplicated rows check the corresponding check box.

To edit a created slot, select it in the *Parameters* area of the *Grouper* element and click the *Edit* button. To remove the slot, select it and click the *Remove* b utton.

Input/Output Ports

The element has 1 *input port* that can take any incoming data.

Name in GUI: *Input data flow*

Name in workflow File: input-data

The element has 1 *output port*.

Name in GUI: *Grouped output data flow*

Name in workflow File: output-data

Slots:

Slot In GUI	Slot in workflow File	Type
Group size	**group-size**	*string*

Also the port has one default slot of the grouped data and it may also have one or several customized output slots (see above).