NiagaraST Operator Functional Specification

|  |  |  |
| --- | --- | --- |
| **Name:** AvgE | **Projects:** Feedback, Windowing | **Author:** RJFM |
| **Revision:** 1 | **Status:** draft | **Date:** 1/25/2010 |

# Preliminaries

An “Early Closure Window” in stream processing is a window that closes before it would normally have, as specified by a closing parameter.

Consider the schema sensors(time, count, speed). A stream s over this schema is punctuated every 5 minutes on attribute time. A query uses 5-minute windows to calculate the average speed of the period. Consider the following tuples:

10:00:00, 25, 50

10:01:00, 45, 39

10:02:00, 20, 45

10:03:00, 14, 44

10:04:00, 22, 42

The average speed for the window [10:00:00 – 10:05:00) is 44.

Now consider an early closing window, which will stop computing the average once at least 30 vehicles have been considered. Such a window could be specified as

<avge id=”avge” groupby=”$wid” avgattr=”$speed” closeattr=”$count” closecondop=”sum” closecond=”30”>

Where one specifies that the sum over the closing attribute is to be monitored, and the aggregate should be output once it reaches (or exceeds) the value 30. In the example stream, we’d see 44.5 as an output.

# The feedback connection

Early closing windows provide opportunities to avoid processing tuples. In this example, three tuples need not be considered for the final output, which can allow us to send feedback to antecedent operators and avoid processing those tuples.

# Implementation considerations

It is incorrect to output another value for an early closed window.