**Шаг 1: Лог событий**

Пример лога событий в формате XES

Trace 1:

- Event: A, Timestamp: 2023-12-01 10:00

- Event: B, Timestamp: 2023-12-01 10:05

- Event: C, Timestamp: 2023-12-01 10:10

Trace 2:

- Event: A, Timestamp: 2023-12-02 11:00

- Event: C, Timestamp: 2023-12-02 11:05

- Event: B, Timestamp: 2023-12-02 11:10

Trace 3:

- Event: B, Timestamp: 2023-12-03 12:00

- Event: C, Timestamp: 2023-12-03 12:05

### ****Шаг 2: Чтение и кодирование событий****

Мы кодируем уникальные события в числовой формат.

* События: A, B, C
* Кодировка:

act\_map = {'A': 0, 'B': 1, 'C': 2, 'start': 3, 'end': 4}

reverse\_map = {0: 'A', 1: 'B', 2: 'C', 3: 'start', 4: 'end'}

### ****Шаг 3: Построение списка пар событий****

Для каждой трассы мы определяем последовательные пары событий:

#### Пример (Trace 1):

Pairs:

(A → B), Timestamp: 2023-12-01 10:05

(B → C), Timestamp: 2023-12-01 10:10

### ****Шаг 4: Агрегация пар в матрицу частот****

Теперь мы разбиваем события на временные интервалы. Допустим:

* **3 временных интервала**:
  1. 2023-12-01 10:00 - 2023-12-01 23:59
  2. 2023-12-02 00:00 - 2023-12-02 23:59
  3. 2023-12-03 00:00 - 2023-12-03 23:59

#### Матрица dfg\_time\_matrix

Мы создаем матрицу [3, 5, 5] (временные интервалы × события × события):

* Переходы между событиями добавляются в матрицу.

Interval 1:

(A → B): +1

(B → C): +1

Interval 2:

(A → C): +1

(C → B): +1

Interval 3:

(B → C): +1

Итоговая матрица:

dfg\_time\_matrix[:, :, :] =

[

[[0, 1, 0, 0, 0], # Interval 1: A → B

[0, 0, 1, 0, 0], # Interval 1: B → C

[0, 0, 0, 0, 0],

[0, 0, 0, 0, 0],

[0, 0, 0, 0, 0]],

[[0, 0, 1, 0, 0], # Interval 2: A → C

[0, 0, 0, 0, 0],

[0, 1, 0, 0, 0], # Interval 2: C → B

[0, 0, 0, 0, 0],

[0, 0, 0, 0, 0]],

[[0, 0, 0, 0, 0],

[0, 0, 1, 0, 0], # Interval 3: B → C

[0, 0, 0, 0, 0],

[0, 0, 0, 0, 0],

[0, 0, 0, 0, 0]],

]

### ****Шаг 5: Подготовка данных для прогнозирования****

Для каждой пары событий (a, a2) мы извлекаем временной ряд переходов:

array = dfg\_time\_matrix[:, a, a2]

#### Пример для пары A → B (a=0, a2=1):

array = [1, 0, 0]

#### Пример для пары B → C (a=1, a2=2):

array = [1, 0, 1]

### ****Шаг 6: Прогнозирование****

* Используем 10-кратную кросс-валидацию (no\_folds = 10).
* Горизонт прогнозирования (horizon = 2).

#### Пример:

Для пары B → C, временной ряд [1, 0, 1]:

* **Fold 1**:
  + Тренировочные данные: [1]
  + Тестовые данные: [0, 1]
  + Прогнозируем горизонтом 2: [y\_hat1, y\_hat2].

#### Итог:

Результаты для каждой модели (dfg\_result\_matrix):

NAV: [1, 1]

ARIMA: [0.8, 1.1]

HW: [0.9, 1.0]