

Esper – CEP

W ramach tego warsztatu chcemy skorzystać z możliwości jakie dają nam:

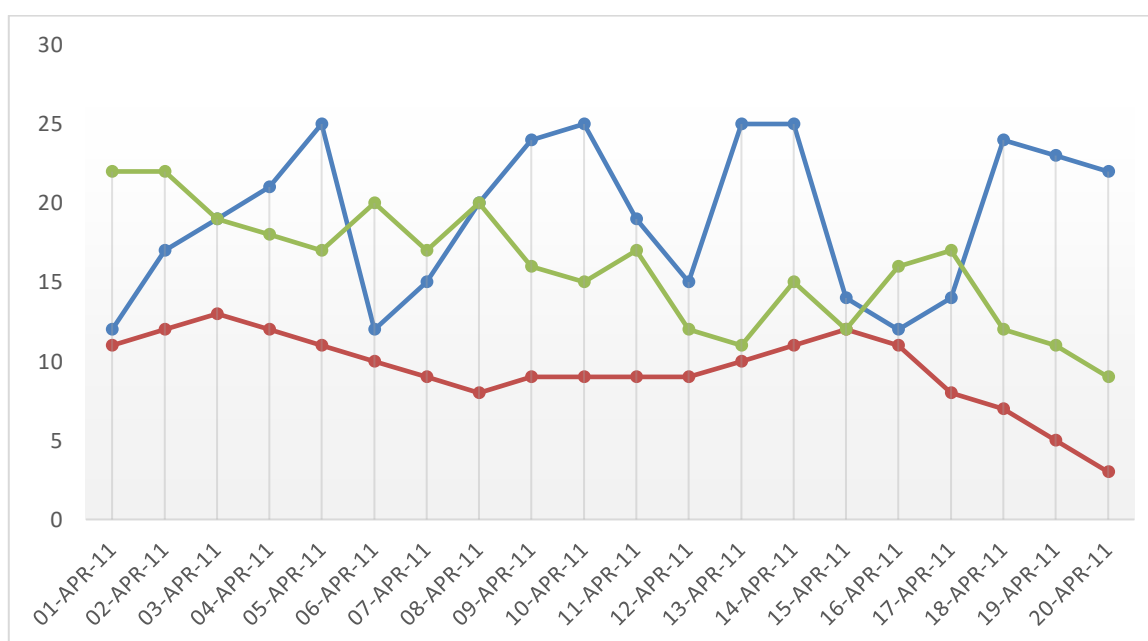
- proste wzorce EPL, a także
- wyrażenia MATCH RECOGNIZE

Do tej pory pracowaliśmy na danych dotyczących rzeczywistych kursów akcji. Moglibyśmy na nich pozostać. Jednak, brak dokładnej ich znajomości, a także spora ilość atrybutów, mogłyby utrudnić nam wykonywanie tych zadań.

W związku z tym skorzystamy z bardzo prostego, statycznego, zbioru danych, który być może już znamy.

https://livesql.oracle.com/apex/livesql/file/tutorial_EG95N4HMCYSL1A0Y6J80UHNHB.html

Dane te zawierają notowania trzech fikcyjnych spółek z przedziału czasu 2011-04-01 - 2011-04-20, a charakterystykę tych danych przedstawia poniższy wykres.



Pobierz wersję początkową naszego projektu

```
git clone https://github.com/BigDataStreamProcessing/esper-CEP.git
```

Przejdź do katalogu projektu, a następnie usuń powiązanie z projektem zewnętrznym

```
cd esper-CEP
git remote rm origin
```

Otwórz katalog esper-CEP w *IntelliJ IDEA*. Uruchom projekt upewniając się, że wszystko działa jak należy.

```
{ "symbol": "OSCOPR", "tstamp": "2011-04-19 00:00:00.0", "price": 11 }
{ "symbol": "ACME", "tstamp": "2011-04-20 00:00:00.0", "price": 22 }
{ "symbol": "GLOBEX", "tstamp": "2011-04-20 00:00:00.0", "price": 3 }
{ "symbol": "OSCOPR", "tstamp": "2011-04-20 00:00:00.0", "price": 9 }

Process finished with exit code 0
```

Wzorce EPL

1. Na początku coś prostego. Zamiast oryginalnego zapytania wstaw poniższe polecenia. Ograniczymy w ten sposób wyszukiwanie wzorców tylko do zdarzeń spółki ACME.

```
create window AcmeTicker#length(10) as Ticker;
insert into AcmeTicker select * from Ticker where symbol = 'ACME';
@name('answer') select a.* from pattern[ a=AcmeTicker(price=25) ]
```

Dlaczego wykrywane jest tylko pierwsze zdarzenie w którym wartość price jest równa 25?

```
{"symbol":"ACME","tstamp":"2011-04-05 00:00:00.0","price":25}
```

Skoryguj zapytanie *answer*, aby wykrywało każde takie zdarzenie.

```
{"symbol":"ACME","tstamp":"2011-04-05 00:00:00.0","price":25}
{"symbol":"ACME","tstamp":"2011-04-10 00:00:00.0","price":25}
{"symbol":"ACME","tstamp":"2011-04-13 00:00:00.0","price":25}
{"symbol":"ACME","tstamp":"2011-04-14 00:00:00.0","price":25}
```

2. Wykryj każdą parę następujących bezpośrednio po sobie notowań spółki ACME. Podaj różnicę pomiędzy cenami tych notowań oraz ich daty.

```
{a.tstamp=2011-04-01 00:00:00.0, diff=5, b.tstamp=2011-04-02 00:00:00.0}
{a.tstamp=2011-04-02 00:00:00.0, diff=2, b.tstamp=2011-04-03 00:00:00.0}
{a.tstamp=2011-04-03 00:00:00.0, diff=2, b.tstamp=2011-04-04 00:00:00.0}
{a.tstamp=2011-04-04 00:00:00.0, diff=4, b.tstamp=2011-04-05 00:00:00.0}
{a.tstamp=2011-04-05 00:00:00.0, diff=-13, b.tstamp=2011-04-06 00:00:00.0}
{a.tstamp=2011-04-06 00:00:00.0, diff=3, b.tstamp=2011-04-07 00:00:00.0}
{a.tstamp=2011-04-07 00:00:00.0, diff=5, b.tstamp=2011-04-08 00:00:00.0}
{a.tstamp=2011-04-08 00:00:00.0, diff=4, b.tstamp=2011-04-09 00:00:00.0}
```

3. Korzystając z warunku umieszczonego w klauzuli WHERE ogranicz te pary tylko do tych, w których nastąpił spadek wartości.

```
{a.tstamp=2011-04-05 00:00:00.0, diff=-13, b.tstamp=2011-04-06 00:00:00.0}
{a.tstamp=2011-04-10 00:00:00.0, diff=-6, b.tstamp=2011-04-11 00:00:00.0}
{a.tstamp=2011-04-11 00:00:00.0, diff=-4, b.tstamp=2011-04-12 00:00:00.0}
{a.tstamp=2011-04-14 00:00:00.0, diff=-11, b.tstamp=2011-04-15 00:00:00.0}
{a.tstamp=2011-04-15 00:00:00.0, diff=-2, b.tstamp=2011-04-16 00:00:00.0}
{a.tstamp=2011-04-18 00:00:00.0, diff=-1, b.tstamp=2011-04-19 00:00:00.0}
{a.tstamp=2011-04-19 00:00:00.0, diff=-1, b.tstamp=2011-04-20 00:00:00.0}
```

4. Znajdź następujące po sobie (niekoniecznie bezpośrednio) pary najbliższych notowań, w których cena przekroczyła 23.

```
{a.tstamp=2011-04-05 00:00:00.0, b.price=24, b.tstamp=2011-04-09 00:00:00.0, a.price=25}
{a.tstamp=2011-04-09 00:00:00.0, b.price=25, b.tstamp=2011-04-10 00:00:00.0, a.price=24}
{a.tstamp=2011-04-10 00:00:00.0, b.price=25, b.tstamp=2011-04-13 00:00:00.0, a.price=25}
{a.tstamp=2011-04-13 00:00:00.0, b.price=25, b.tstamp=2011-04-14 00:00:00.0, a.price=25}
{a.tstamp=2011-04-14 00:00:00.0, b.price=24, b.tstamp=2011-04-18 00:00:00.0, a.price=25}
```

5. Powyższe pary tworzą interwały, w których spółka osiągnęła niższe wartości. Ogranicz te interwały do takich, w czasie których cena spółki nie spadła poniżej 14.

```
{a.timestamp=2011-04-09 00:00:00.0, b.price=25, b.timestamp=2011-04-10 00:00:00.0, a.price=24}
{a.timestamp=2011-04-10 00:00:00.0, b.price=25, b.timestamp=2011-04-13 00:00:00.0, a.price=25}
{a.timestamp=2011-04-13 00:00:00.0, b.price=25, b.timestamp=2011-04-14 00:00:00.0, a.price=25}
```

6. Uzyskuj informacje o seriach notowań mających długość co najmniej 3, w których wartość ceny nie spadała poniżej 20. Pobieraj wartości ceny dla pierwszych trzech notowań oraz datę rozpoczęcia serii.

```
{p1=20, tstart=2011-04-08 00:00:00.0, p2=24, p3=25}
```

Wyrażenie MATCH_RECOGNIZE

Kolejne zadania rozwiąż korzystając z wyrażenia MATCH_RECOGNIZE (MR)

7. Na początku klasyka. Wykrywaj serie notowań układających się w kształt litery V. Zwróć uwagę na to, że seria nie musi składać się z trzech zdarzeń. Pamiętaj o ograniczeniach w wyrażeniu MR w systemie Esper. W razie potrzeby zaglądaj do dokumentacji

<https://esper.espertech.com/release-8.9.0/reference-esper/html/match-recognize.html>

```
{startPrice=25, symbol=ACME, lowPrice=12, highPrice=15, stopT=2011-04-07 00:00:00.0, startT=2011-04-05 00:00:00.0}
{startPrice=25, symbol=ACME, lowPrice=15, highPrice=25, stopT=2011-04-13 00:00:00.0, startT=2011-04-10 00:00:00.0}
{startPrice=25, symbol=ACME, lowPrice=12, highPrice=14, stopT=2011-04-17 00:00:00.0, startT=2011-04-14 00:00:00.0}
```

Przeglądnij się dokładnie Twojemu rozwiązaniu i porównaj go z uzyskanym powyżej wynikiem.

Jeśli Twoje daty nie w pełni obejmują kształty V, zastanów się co może być przyczyną.

8. Przestań analizować tylko spółkę ACME. Zmień źródło na wszystkie zdarzenia ze schematu Ticker. Ponownie uruchom swoje przetwarzanie wykrywające serie notowań układających się w kształt litery V. Pamiętaj, że wyrażenie MR musi działać na strumieniu ograniczonym za pomocą okna.

```
{startPrice=19, symbol=OSCORP, lowPrice=17, highPrice=20, stopT=2011-04-06 00:00:00.0, startT=2011-04-03 00:00:00.0}
{startPrice=25, symbol=ACME, lowPrice=12, highPrice=15, stopT=2011-04-07 00:00:00.0, startT=2011-04-05 00:00:00.0}
{startPrice=10, symbol=GLOBEX, lowPrice=8, highPrice=9, stopT=2011-04-09 00:00:00.0, startT=2011-04-06 00:00:00.0}
{startPrice=25, symbol=ACME, lowPrice=15, highPrice=25, stopT=2011-04-13 00:00:00.0, startT=2011-04-10 00:00:00.0}
{startPrice=15, symbol=OSCORP, lowPrice=11, highPrice=15, stopT=2011-04-14 00:00:00.0, startT=2011-04-11 00:00:00.0}
{startPrice=25, symbol=ACME, lowPrice=12, highPrice=14, stopT=2011-04-17 00:00:00.0, startT=2011-04-14 00:00:00.0}
```

Jeśli chcesz możesz sprawdzić czy analogiczne wyniki otrzymasz na platformie udostępnianej przez

EsperTech - <http://esper-epl-tryout.appspot.com/epltryout/mainform.html>

- a. Do pola *Enter EPL Here*: wprowadź polecenie tworzące nasz schemat zdarzeń (podane poniżej), a następnie po adnotacji @Name('Out') umieść Twoje rozwiązanie (zapytanie)

```
create schema Ticker(symbol string, tstamp string, price int);
```

```
@Name( 'Out' )
```


b. Do pola *Enter sequence of time and events*: wprowadź

```

Ticker={symbol='ACME', tstamp='2011-04-01 00:00:00.0', price=12}
Ticker={symbol='GLOBEX', tstamp='2011-04-01 00:00:00.0', price=11}
Ticker={symbol='OSCORP', tstamp='2011-04-01 00:00:00.0', price=22}
Ticker={symbol='ACME', tstamp='2011-04-02 00:00:00.0', price=17}
Ticker={symbol='GLOBEX', tstamp='2011-04-02 00:00:00.0', price=12}
Ticker={symbol='OSCORP', tstamp='2011-04-02 00:00:00.0', price=22}
Ticker={symbol='ACME', tstamp='2011-04-03 00:00:00.0', price=19}
Ticker={symbol='GLOBEX', tstamp='2011-04-03 00:00:00.0', price=13}
Ticker={symbol='OSCORP', tstamp='2011-04-03 00:00:00.0', price=19}
Ticker={symbol='ACME', tstamp='2011-04-04 00:00:00.0', price=21}
Ticker={symbol='GLOBEX', tstamp='2011-04-04 00:00:00.0', price=12}
Ticker={symbol='OSCORP', tstamp='2011-04-04 00:00:00.0', price=18}
Ticker={symbol='ACME', tstamp='2011-04-05 00:00:00.0', price=25}
Ticker={symbol='GLOBEX', tstamp='2011-04-05 00:00:00.0', price=11}
Ticker={symbol='OSCORP', tstamp='2011-04-05 00:00:00.0', price=17}
Ticker={symbol='ACME', tstamp='2011-04-06 00:00:00.0', price=12}
Ticker={symbol='GLOBEX', tstamp='2011-04-06 00:00:00.0', price=10}
Ticker={symbol='OSCORP', tstamp='2011-04-06 00:00:00.0', price=20}
Ticker={symbol='ACME', tstamp='2011-04-07 00:00:00.0', price=15}
Ticker={symbol='GLOBEX', tstamp='2011-04-07 00:00:00.0', price=9}
Ticker={symbol='OSCORP', tstamp='2011-04-07 00:00:00.0', price=17}
Ticker={symbol='ACME', tstamp='2011-04-08 00:00:00.0', price=20}
Ticker={symbol='GLOBEX', tstamp='2011-04-08 00:00:00.0', price=8}
Ticker={symbol='OSCORP', tstamp='2011-04-08 00:00:00.0', price=20}
Ticker={symbol='ACME', tstamp='2011-04-09 00:00:00.0', price=24}
Ticker={symbol='GLOBEX', tstamp='2011-04-09 00:00:00.0', price=9}
Ticker={symbol='OSCORP', tstamp='2011-04-09 00:00:00.0', price=16}
Ticker={symbol='ACME', tstamp='2011-04-10 00:00:00.0', price=25}
Ticker={symbol='GLOBEX', tstamp='2011-04-10 00:00:00.0', price=9}
Ticker={symbol='OSCORP', tstamp='2011-04-10 00:00:00.0', price=15}
Ticker={symbol='ACME', tstamp='2011-04-11 00:00:00.0', price=19}
Ticker={symbol='GLOBEX', tstamp='2011-04-11 00:00:00.0', price=9}
Ticker={symbol='OSCORP', tstamp='2011-04-11 00:00:00.0', price=15}
Ticker={symbol='ACME', tstamp='2011-04-12 00:00:00.0', price=15}
Ticker={symbol='GLOBEX', tstamp='2011-04-12 00:00:00.0', price=9}
Ticker={symbol='OSCORP', tstamp='2011-04-12 00:00:00.0', price=12}
Ticker={symbol='ACME', tstamp='2011-04-13 00:00:00.0', price=25}
Ticker={symbol='GLOBEX', tstamp='2011-04-13 00:00:00.0', price=10}
Ticker={symbol='OSCORP', tstamp='2011-04-13 00:00:00.0', price=11}
Ticker={symbol='ACME', tstamp='2011-04-14 00:00:00.0', price=25}
Ticker={symbol='GLOBEX', tstamp='2011-04-14 00:00:00.0', price=11}
Ticker={symbol='OSCORP', tstamp='2011-04-14 00:00:00.0', price=15}
Ticker={symbol='ACME', tstamp='2011-04-15 00:00:00.0', price=14}
Ticker={symbol='GLOBEX', tstamp='2011-04-15 00:00:00.0', price=12}
Ticker={symbol='OSCORP', tstamp='2011-04-15 00:00:00.0', price=12}
Ticker={symbol='ACME', tstamp='2011-04-16 00:00:00.0', price=12}
Ticker={symbol='GLOBEX', tstamp='2011-04-16 00:00:00.0', price=11}
Ticker={symbol='OSCORP', tstamp='2011-04-16 00:00:00.0', price=16}
Ticker={symbol='ACME', tstamp='2011-04-17 00:00:00.0', price=14}
Ticker={symbol='GLOBEX', tstamp='2011-04-17 00:00:00.0', price=8}
Ticker={symbol='OSCORP', tstamp='2011-04-17 00:00:00.0', price=14}
Ticker={symbol='ACME', tstamp='2011-04-18 00:00:00.0', price=24}
Ticker={symbol='GLOBEX', tstamp='2011-04-18 00:00:00.0', price=7}
Ticker={symbol='OSCORP', tstamp='2011-04-18 00:00:00.0', price=12}
Ticker={symbol='ACME', tstamp='2011-04-19 00:00:00.0', price=23}
Ticker={symbol='GLOBEX', tstamp='2011-04-19 00:00:00.0', price=5}
Ticker={symbol='OSCORP', tstamp='2011-04-19 00:00:00.0', price=11}
Ticker={symbol='ACME', tstamp='2011-04-20 00:00:00.0', price=22}
Ticker={symbol='GLOBEX', tstamp='2011-04-20 00:00:00.0', price=3}
Ticker={symbol='OSCORP', tstamp='2011-04-20 00:00:00.0', price=9}

```

- c. Skorzystaj z przycisku *Submit*, aby uruchomić przetwarzanie



Esper EPL Online

Esper Version 8.8.0

[Terms of use](#)
[Use Esper Notebook](#)

[Help](#)
[Reset Form](#)
[Clear Form](#)

EPL Statements

EPL Module Text

Enter EPL Here:

```
create schema Ticker(symbol string, tstamp string, price int);

@Name('Out') select *
from Ticker#length(10)
match_recognize (
partition by symbol
measures
strt.symbol as symbol,
strt.tstamp as startT,
```

Time And Event Sequence

Beginning Of Time

Provide a timestamp to start at:

2001-01-01 08:00:00.000

Submit

Advance Time and Send Events

Enter sequence of time and events:

```
Ticker=(symbol='ACME', tstamp='2011-04-10 00:00:00.0', price=23)
Ticker=(symbol='GLOBEX', tstamp='2011-04-10 00:00:00.0', price=9)
Ticker=(symbol='OSCORP', tstamp='2011-04-10 00:00:00.0', price=15)
Ticker=(symbol='ACME', tstamp='2011-04-11 00:00:00.0', price=19)
Ticker=(symbol='GLOBEX', tstamp='2011-04-11 00:00:00.0', price=9)
Ticker=(symbol='OSCORP', tstamp='2011-04-11 00:00:00.0', price=15)
```

Scenario Results

All Output Events

Output Per Statement

All Audit Text

Audit Text Per Statement

At: 2001-01-01 08:00:00.000

Statement: Out

Insert

```
stmt1_mrow2=
{symbol='OSCORP',
startT='2011-04-03
```

Aby zauważyć różnice pomiędzy wyrażeniem MR w funkcjonalności oferowanej przez *Esper*, a implementacją w bazie danych *Oracle* warto spróbować uzyskać ten sam efekt, na tych samych danych korzystając z tutorialu

https://livesql.oracle.com/apex/livesql/file/tutorial1_EG95N4HMCYSL1A0Y6J80UHNHB.html

Uwaga! Konieczne jest konto (darmowe) na platformie Oracle.

Na zakończenie, dwa proste zadania.

- Analizując każdą ze spółek niezależnie, znajdź serię dwóch kolejnych podwyżek ceny takich, że druga podwyżka była ponad dwa razy większa niż pierwsza. W których systemach (Esper, Oracle) da się to rozwiązać całkowicie w ramach wyrażenia MR?
- No i może powrót do klasyki – litera W. Nie trać zbyt dużo czasu nad rozwiązaniem tego zadania przy użyciu MR... no chyba, że udało Ci się sprawić, aby Twoje litery V były zachłanne (*greedy*). Tam gdzie, w przypadku Espera, MR nie może, tam może wzorzec EPL da radę? Podejmujesz wyzwanie?

SYMBOL	START_TS	BOTTOM_TS	END_TS
ACME	06-APR-11	12-APR-11	13-APR-11
OSCORP	03-APR-11	07-APR-11	08-APR-11
OSCORP	12-APR-11	15-APR-11	16-APR-11

