

ChatGPT's conversation

Let \mathcal{E} denote an event, then:

$$\mathcal{E} = \langle \mathcal{F}, \mathcal{Y}, \mathcal{D}, \mathcal{A}, \mathcal{C}, \mathcal{R} \rangle$$

where:

- \mathcal{F} is the set of affected firms by the event. Formally, $\mathcal{F} \subseteq \mathcal{S}$, where \mathcal{S} denotes the universal set of all firms being considered in the analysis.
- \mathcal{Y} is the datetime at which the event occurs. This could be represented as a specific timestamp in the appropriate format, typically in ISO 8601 format: $\mathcal{Y} \in \mathbb{T}$, where \mathbb{T} denotes the set of all possible datetimes.
- \mathcal{D} represents the description of the event. This could be a natural language text summary or a vector representation (e.g., an embedding in some vector space) that captures the semantic meaning of the event.
- \mathcal{A} represents the attributes of the event, which could include various quantitative and qualitative measures, such as the magnitude of the event, sentiment scores, and other features derived from the news articles. Formally, $\mathcal{A} = \{a_1, a_2, \dots, a_k\}$, where each $a_i \in \mathbb{R}$ represents a measurable attribute.
- \mathcal{C} is the category or type of event. For example, this could be an economic event, a merger/acquisition announcement, a regulatory change, etc. Formally, \mathcal{C} belongs to a predefined set of categories, $\mathcal{C} \in \mathcal{K}$, where \mathcal{K} is a finite set of possible event types.
- \mathcal{R} is the relevance or impact measure of the event, which quantifies how significant the event is with respect to the firms in \mathcal{F} . This could be a vector or a single value depending on whether you wish to capture the impact per firm or globally. Formally, \mathcal{R} could be a mapping $\mathcal{R} : \mathcal{F} \rightarrow \mathbb{R}$, assigning an impact score to each affected firm, or a single value $r \in \mathbb{R}$.

Objective Event \mathcal{E}^o

An Objective Event \mathcal{E}^o is an unobservable event that has actually occurred. It is characterized as:

$$\mathcal{E}^o = \langle \mathcal{F}, \mathcal{Y}, \mathcal{D}^o, \mathcal{A}^o, \mathcal{C}, \mathcal{R}^o \rangle$$

where all components are as defined previously. However, the investor does not have direct access to \mathcal{E}^o .

Subjective Event \mathcal{E}_i^s

A Subjective Event \mathcal{E}_i^s is the event as perceived and constructed by the investor after reading news articles. It is denoted as:

$$\mathcal{E}_i^s = \langle \mathcal{F}_i^s, \mathcal{Y}_i^s, \mathcal{D}_i^s, \mathcal{A}_i^s, \mathcal{C}_i^s, \mathcal{R}_i^s \rangle$$

where:

- \mathcal{F}_i^s is the set of firms the investor believes are affected, derived from the information in the articles.
- \mathcal{Y}_i^s is the datetime the investor believes the event took place, based on the articles.
- \mathcal{D}_i^s is the investor's subjective description of the event, formed after synthesizing the information.
- \mathcal{A}_i^s represents the attributes (e.g., magnitude, sentiment) that the investor assigns to the event.
- \mathcal{C}_i^s is the category of the event as perceived by the investor.
- \mathcal{R}_i^s is the perceived relevance or impact of the event on the market, according to the investor's interpretation.

Information Gathering and Event Formation

The investor gathers information by reading news articles. Let $\mathcal{N}_{i,j}$ represent the j -th news article about event i . The investor reads n_i articles about event i from different sources, forming a set:

$$\mathcal{N}_i = \{\mathcal{N}_{i,1}, \mathcal{N}_{i,2}, \dots, \mathcal{N}_{i,n_i}\}$$

Each article $\mathcal{N}_{i,j}$ provides a reported event $\mathcal{E}_{i,j}^r$, which the investor uses to infer the subjective event \mathcal{E}_i^s . The formation of the subjective event can be formalized as:

$$\mathcal{E}_i^s = f\left(\left\{\mathcal{E}_{i,j}^r\right\}_{j=1}^{n_i}\right)$$

where f is a function that models the investor's process of aggregating and synthesizing the information from the various news articles. This function could involve: - Weighting the importance of different sources. - Averaging or voting on common attributes. - Applying heuristics based on the investor's experience or biases.

Trading Signal Formation

After forming the subjective event \mathcal{E}_i^s , the investor assigns a trading signal S_i based on his perception of the event. The trading signal could be a buy, sell, or hold decision, represented as:

$$S_i = g(\mathcal{E}_i^s)$$

where g is the decision function, which might depend on: - The perceived impact \mathcal{R}_i^s . - The investor's risk appetite. - The investor's overall strategy (e.g., momentum, mean-reversion).

Trading Strategy Based on Subjective Events

The investor executes trades in the market based on the trading signals S_i generated by the subjective events \mathcal{E}_i^s . Let \mathcal{T} denote the set of trades:

$$\mathcal{T} = \{T_i\}_{i=1}^m$$

where each trade T_i is executed based on the trading signal S_i . The outcome of these trades depends on how well the subjective events \mathcal{E}_i^s approximate the objective events \mathcal{E}^o .

Objective vs. Subjective Event Dynamics

The key dynamic here is the gap between the subjective events \mathcal{E}_i^s and the unobservable objective events \mathcal{E}^o . This gap introduces a level of uncertainty and risk in the investor's trading strategy.

- **Modeling Errors:** If the subjective event \mathcal{E}_i^s diverges significantly from \mathcal{E}^o , the investor's trades might lead to losses. This can be modeled by comparing the performance of trades \mathcal{T} with the theoretical performance had the investor known \mathcal{E}^o .
- **Learning and Adaptation:** Over time, the investor might adapt f and g based on the outcomes of their trades, refining how they infer subjective events and assign trading signals.

Given the uncertainty about the true nature of \mathcal{E}^o , the investor might model the subjective events probabilistically:

$$\mathcal{E}_i^s \sim P\left(\mathcal{E}_i^o \mid \left\{\mathcal{E}_{i,j}^r\right\}_{j=1}^{n_i}\right)$$

Formalization in a Probabilistic Framework

where $P\left(\mathcal{E}_i^o \mid \left\{\mathcal{E}_{i,j}^r\right\}_{j=1}^{n_i}\right)$ represents the posterior distribution of the objective event given the reported events. The investor's trading decision can then be viewed as an action taken under uncertainty, with the goal of maximizing expected utility:

$$S_i = \arg \max_{a \in \text{Actions}} \mathbb{E}[U(a \mid \mathcal{E}_i^s)]$$