From inference to influence: applying causal game theory to complex security environments

Vonk, M.C.

Chapter 2

Preliminaries

This chapter introduces preliminaries about probability theory, graph theory, and game theory. Frequently used acronyms and notation conventions are introduced in Appendix A.

2.1 Probability Theory

This section discusses preliminaries of probability theory and the notation conventions followed throughout the dissertation. Probability theory revolves around events to which probability can be ascribed.

Definition 2.1 (State Space)**.** The *state space* Ω is the set of possible outcomes of an event.

Definition 2.2 (Event)**.** An *event* is a subset of Ω to which probability can be ascribed. The collection of all events is denoted by $\mathcal{F}$.

Definition 2.3 (Probability Measure)**.** A *probability measure* is a function $P : \mathcal{F} \to [0, 1]$ that satisfies the following requirements:

- $P(\Omega) = 1$.
- For all countable disjoint $C_1, C_2, \ldots \in \mathcal{F}$, the following holds:

$$P\left(\bigcup_{i=1}^{\infty} C_i\right) = \sum_{i=1}^{\infty} P(C_i).$$

Definition 2.4 (Probability Space)**.** The triplet $(\Omega, \mathcal{F}, P)$ is called the *probability space.*

Definition 2.5 (Conditional Probability)**.** For $C_1, C_2 \in \mathcal{F}$ and $P(C_2) > 0$, the *conditional probability* of C_1 given C_2 is defined as:

$$P(C_1 \mid C_2) = \frac{P(C_1 \cap C_2)}{P(C_2)}.$$

Definition 2.6 (Chain Rule)**.** Let $C_1, C_2 \in \mathcal{F}$ with $P(C_2) > 0$. The *chain rule* of probability expresses the joint probability of two events as:

$$P(C_1 \cap C_2) = P(C_2)P(C_1 \mid C_2).$$

Definition 2.7 (Bayes' Rule)**.** For $C_1, C_2 \in \mathcal{F}$ and $P(C_2) > 0$, *Bayes' rule* relates conditional probabilities as:

$$P(C_1 \mid C_2) = \frac{P(C_2 \mid C_1)P(C_1)}{P(C_2)}.$$

The concepts of random variables and their probability distributions provide a structured way to model and quantify uncertainty in real-world phenomena.

Definition 2.8 (Random Variable)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space. A *random variable* X is a function $X : \Omega \to \mathbb{R}$ such that for every $B \subset \mathbb{R}$:[1]

$$X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathcal{F}.$$

Definition 2.9 (Probability Distribution)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $X : \Omega \to \mathbb{R}$ be a random variable. The *probability distribution* of X is the probability measure P induced by X, which assigns probabilities to subsets of $\mathbb{R}$, and is defined as:

$$P(B) = P(X^{-1}(B)) = P(\{\omega \in \Omega \mid X(\omega) \in B\}), \quad \forall B \subset \mathbb{R}.$$

Random variables are denoted by capital letters and $\boldsymbol{X} = \{X_1, \ldots, X_n\}$ denotes the set containing random variables X_i that take values x_i in the corresponding state space Ω_{X_i}. The probability that random variable X_i takes value x_i is denoted by

[1] Technically, this condition should hold for every B in the Borel σ-algebra on $\mathbb{R}$, not all subsets. A full treatment of Borel sets is beyond the scope of this dissertation.

$P(X_i = x_i)$ or, in short, $P(x_i)$. If X_i is discrete, the probability distribution is denoted by uppercase $P(X_i)$. When X_i is continuous, it is denoted by lowercase $p(X_i)$.

Definition 2.10 (Joint Distribution)**.** The *joint distribution* of random variables $X_1, \ldots, X_n$ describes the probability distribution of their simultaneous occurrences and is denoted by $P(X_1, \ldots, X_n)$ or $P(\boldsymbol{X})$.

Definition 2.11 (Marginal Distribution)**.** A *marginal distribution* of a single random variable X_i taking value x_i is obtained by *marginalizing* the other random variables $\boldsymbol{X}_{-i}$ out:

- $P(x_i) = \sum_{\boldsymbol{x}_{-i} \in \Omega_{\boldsymbol{X}_{-i}}} P(x_i, \boldsymbol{x}_{-i})$ when the random variables are discrete.
- $p(x_i) = \int_{\Omega_{\boldsymbol{X}_{-i}}} p(x_i, \boldsymbol{x}_{-i}) d\boldsymbol{x}_{-i}$ when the random variables are continuous.

Definition 2.12 (Expected Value)**.** The *expected value* of a discrete random variable X_i under probability distribution P is

$$\mathbb{E}[X_i] = \sum_{x_i \in \Omega_{X_i}} x_i P(x_i)$$

while the *expected value* of a continuous random variable X_i under probability distribution p is

$$\mathbb{E}[X_i] = \int_{\Omega_{X_i}} x_i p(x_i) dx_i.$$

2.2 Graphical Models

The following graph-theoretic definitions underpin the causal inference concepts used throughout this thesis.

A graph is denoted by $G = (\boldsymbol{V}, \boldsymbol{E})$, where $\boldsymbol{V} = \{V_1, \ldots, V_n\}$ is the set of vertices (or nodes) and $\boldsymbol{E} \subseteq \boldsymbol{V} \times \boldsymbol{V}$ is the set of edges.

Definition 2.13 (Directed, Undirected, and Partially Directed Graphs)**.** A graph $G = (\boldsymbol{V}, \boldsymbol{E})$ is called:

- *directed* if every edge $(V_i, V_j) \in \boldsymbol{E}$ has an assigned direction from V_i to V_j;
- *undirected* if no edge in $\boldsymbol{E}$ has a direction;
- *partially directed* if $\boldsymbol{E}$ contains both directed and undirected edges.

For directed graphs, a fundamental structural property is the presence or absence of cycles.

Definition 2.14 (Cycle)**.** A *cycle* in a directed graph is a sequence of vertices $V_{i_1}, \dots, V_{i_k}$ where $k \geq 2$, $(V_{i_j}, V_{i_{j+1}}) \in \boldsymbol{E}$ for all $j = 1, \dots, k-1$, and $(V_{i_k}, V_{i_1}) \in \boldsymbol{E}$.

Definition 2.15 (Directed Acyclic Graph)**.** A *directed acyclic graph* (DAG) is a directed graph that contains no cycles.

In addition to standard directed and undirected edges, bidirected edges are used to represent certain causal structures.

Definition 2.16 (Bidirected Edge)**.** A *bidirected edge* between vertices V_i and V_j is denoted by $V_i \leftrightarrow V_j$ and represents a specific edge type distinct from both directed edges ($V_i \to V_j$) and undirected edges ($V_i - V_j$).

Definition 2.17 (Acyclic Directed Mixed Graph)**.** An *acyclic directed mixed graph* (ADMG) is a graph $G = (\boldsymbol{V}, \boldsymbol{E})$ where $\boldsymbol{E}$ consists only of directed and bidirected edges, and the directed edges form no cycles.

The following relational concepts describe how vertices are connected within directed graphs.

Definition 2.18 (Parent and Child)**.** Let $G = (\boldsymbol{V}, \boldsymbol{E})$ be a directed graph. If $(V_1, V_2) \in \boldsymbol{E}$ (denoted $V_1 \to V_2$), then V_1 is called a *parent* of V_2 and V_2 is called a *child* of V_1. The set of parents of V_i is denoted by $\mathbf{pa}(V_i) = \{V_j \in \boldsymbol{V} \mid (V_j, V_i) \in \boldsymbol{E}\}$, and the set of children of V_i by $\mathbf{ch}(V_i) = \{V_j \in \boldsymbol{V} \mid (V_i, V_j) \in \boldsymbol{E}\}$.

Definition 2.19 (Ancestor and Descendant)**.** Let $G = (\boldsymbol{V}, \boldsymbol{E})$ be a directed graph. An *ancestor* of node V_i is any node V_j for which there exists a directed path from V_j to V_i, including V_i itself. The set of ancestors is denoted by $\mathbf{an}(V_i)$. Similarly, a *descendant* of V_i is any node V_j for which there exists a directed path from V_i to V_j, including V_i itself. The set of descendants is denoted by $\mathbf{de}(V_i)$.

Definition 2.20 (Non-descendants)**.** The set of *non-descendants* of V_i is defined as $\mathbf{nonde}(V_i) = \boldsymbol{V} \setminus \mathbf{de}(V_i)$. Note that $V_i \notin \mathbf{nonde}(V_i)$ since $V_i \in \mathbf{de}(V_i)$.

For acyclic graphs, a linear ordering can be imposed on the nodes that respects the edge directions.

Definition 2.21 (Topological Sort)**.** A *topological sort* is a total ordering $<$ on $\boldsymbol{V}$ such that $(V_i, V_j) \in \boldsymbol{E}$ implies $V_i < V_j$. A topological sort exists if and only if the graph is acyclic.

Two important graph operations involve restricting the vertex set (subgraphs) or removing specific edges (mutilated graphs).

Definition 2.22 (Subgraph)**.** Given a topological sort $<$ on $\boldsymbol{V}$, the *subgraph* G_i is the induced subgraph on the vertex set $\{V_j \in \boldsymbol{V} \mid V_j \leq V_i\}$, i.e., the graph restricted to nodes that precede or equal V_i in the topological ordering.

Definition 2.23 (Mutilated Graph)**.** Given a subset $\boldsymbol{V}' \subseteq \boldsymbol{V}$, the *mutilated graph* $G_{\overline{\boldsymbol{V}'}}$ is the graph $(\boldsymbol{V}, \boldsymbol{E}')$ where $\boldsymbol{E}' = \boldsymbol{E} \setminus \{(V_i, V_j) \in \boldsymbol{E} \mid V_j \in \boldsymbol{V}'\}$, i.e., all edges directed into nodes in $\boldsymbol{V}'$ are removed.

When graphs are endowed with probabilistic meaning, random variables $\boldsymbol{X} = \{X_1, \ldots, X_n\}$ will correspond to nodes of the graph $\boldsymbol{V} = \{V_1, \ldots, V_n\}$ and therefore $\boldsymbol{V}$ will inherit the probability distributions and state spaces from $\boldsymbol{X}$ (meaning $P(\boldsymbol{V})$ and v_i will correspond to $P(\boldsymbol{X})$ and x_i, respectively). In this case, $\mathbf{pa}(V_i)$ refers to the random variables that are associated with the parents of V_i. The assignment of random variables $\mathbf{pa}(V_i)$ is denoted by $\boldsymbol{pa}_i$, which is an element of state space $\Omega_{\mathbf{pa}(V_i)}$.

2.3 Game Theory

A game models strategic interactions among rational decision-makers, called *agents* or *players*, where each decision-maker selects actions to maximize their own objectives while considering the choices of others. It provides a structured framework to analyze decision-making in competitive or cooperative settings.

Definition 2.24 (Game)**.** A *game* is denoted by $\Gamma = (M, \boldsymbol{A}, \boldsymbol{U})$ and consists of:

- A finite set of *players* $M = \{1, \ldots, m\}$.
- A set of *action sets* $\boldsymbol{A} = \{A^1, \ldots, A^m\}$, where A^i is the set of available actions for player $i \in M$.
- A set of *utility functions* $\boldsymbol{U} = \{u^1, \ldots, u^m\}$, where $u^i : \prod_{j \in M} A^j \to \mathbb{R}$ maps each action profile $(a^1, \ldots, a^m)$ with $a^j \in A^j$ to a real-valued payoff for player i.

An *action profile* is a tuple $a = (a^1, \ldots, a^m) \in \prod_{j \in M} A^j$ specifying one action for each player. For a given player i, the notation $a^{-i} = (a^1, \ldots, a^{i-1}, a^{i+1}, \ldots, a^m)$ denotes the *partial action profile* of all players except i, so that $a = (a^i, a^{-i})$.

Definition 2.25 (Strategy)**.** A *strategy* (or *mixed strategy*) for player $i \in M$ is a probability distribution $\sigma^i \in \Delta(A^i)$, where $\Delta(A^i)$ denotes the simplex over the action set A^i. That is, σ^i assigns a probability $\sigma^i(a^i) \geq 0$ to each action $a^i \in A^i$ such that $\sum_{a^i \in A^i} \sigma^i(a^i) = 1$. The set of all strategies for player i is denoted by $\mathbf{\Sigma}^i = \Delta(A^i)$. A strategy is called *pure* if $\sigma^i(a^i) \in \{0, 1\}$ for all $a^i \in A^i$, meaning the player deterministically selects a single action, and *fully mixed* if $\sigma^i(a^i) > 0$ for all $a^i \in A^i$, meaning every action is chosen with positive probability.

A *strategy profile* $\sigma = (\sigma^1, \ldots, \sigma^m)$ specifies a strategy for each player. The expected utility for player i under a strategy profile σ is denoted $u^i(\sigma)$.

Definition 2.26 (Strategy Profile)**.** A *strategy profile* is a tuple $\sigma = (\sigma^1, \ldots, \sigma^m) \in \prod_{j \in M} \mathbf{\Sigma}^j$ specifying one strategy for each player. For a given player i, the notation $\sigma^{-i} = (\sigma^1, \ldots, \sigma^{i-1}, \sigma^{i+1}, \ldots, \sigma^m)$ denotes the *partial strategy profile* of all players except i, so that $\sigma = (\sigma^i, \sigma^{-i})$.