Signed Graphs and Applications to Social Science

Connor Stitt

MA 351 (001)

14 November 2024

"The enemy of my enemy is my friend."

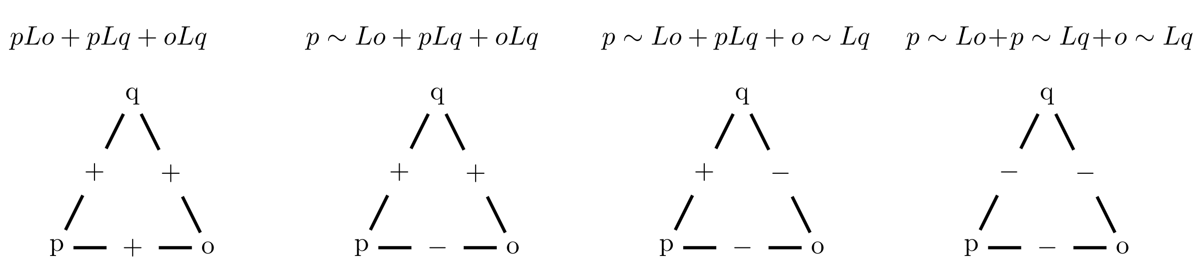
Introduction and the Theory of Structural Balance

In the context of building mathematical models, graphs are useful abstractions that empower researchers to describe relational phenomena. Moreover, graphs have a dual character in the sense that they have both intuitive visual representations and useful matrix forms. However, models of certain real-life phenomena may require additional structure in order to account for their complexity. Such structure can be achieved by modifying existing graphs with additional features, such as adding a scalar weight or a sign. Chiefly, this paper will examine signed graphs will outline the basic properties of the cycle basis and its role in verifying balance between groups in a social networks. Furthermore, this paper will investigate the applications of signed graphs in social science, particularly in mapping coalition networks in Congress and in describing resource allocation in international relations.

Recall that a graph G has a vertex set V and an edge set E. Then, a signed graph is one whose edges are each affixed with either a positive or negative sign, which may symbolize sympathetic or antagonistic relationships, respectively. Put more formally, a signed graph $G(V, E, \sigma)$ is a graph $G = (V, E)$ together with a function $\sigma : E \to \{+, -\}$ which attaches a sign to each edge [2]. The signs therefore represent strong, albeit unitless, symmetric relationships between any two vertices v and u in V [7]. In this sense, if v and u share a positive edge, they are friends; if v and u share a negative edge, they are enemies.

Signed graphs first arose in the field of sociology. Although first published with a logical rather than graph theoretic construction, psychologist Fritz Heider's 1946 study of interpersonal attitudes in small groups opened the pathway to a strict analysis of group dynamics [7]. Heider was interested in how local information, such as the relationships between two individuals, impacted the global phenomenon of balance, particularly in triads of individuals. In his article *Attitudes and Cognitive Organization*, Heider argues that a "balanced configuration exists if the attitudes towards the parts of a causal unit are similar." [4] In this study, the "causal unit" is a triad where the elements are also allowed to include events, items and entities in addition to people. Put simply, Heider wanted to outline which exact scenarios minimized tension between the agents of the causal unit.

In Heider's logical construction, for three people p, o and q, a positive relation between two people is denoted L, whereas a negative relation is denoted $\sim L$. Between any two people and an outside entity, such as a political position, a dissonance may arise when two friends disagree. That disagreement, represented by figure (b), can be resolved if one of the friends change their opinion. A decade later, two mathematicians, Cartwright and Harary, extended Heider's theory of balance to graphs. What they found was that all imbalanced triads included an odd number of negative signs [7]. One could equivalently multiply the signs of each cycle to verify if a graph is balanced. If all cycles have a positive product, then the graph is balanced. Below are the Heider's logically stated relationships for triads, and their graph counterparts.



(a) Balanced Graph (b) Unbalanced Graph (c) Balanced Graph (d) Unbalanced Graph

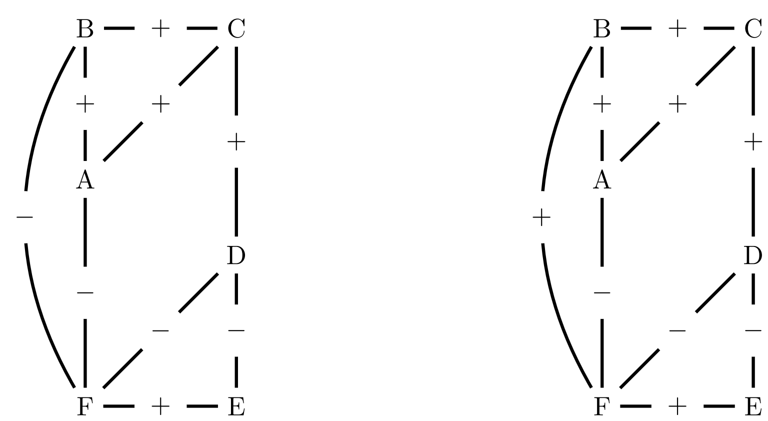
Figure 1: Heider's Triads

Balance in a signed graph could heuristically be described as an "absence of tension" between different actors, such as nations or policymakers, where they work well together [7]. Moreover, these triads are too limited to describe large social networks. Harary's structure theorem generalizes the notion of balance from 3-cycles to cyclic graphs of any size.

Theorem 1 (Structure Theorem) [7]:
Suppose G is a signed graph. Then the following are equivalent:

1. G is balanced.
2. Every closed chain in G is positive.
3. Any two chains between vertices u and v have the same sign.
4. The set V can be partitioned into two sets A and B such that every positive edge joins vertices of the same set and every negative edge joins vertices of different sets.

The primary interpretation of this theorem is that within any balanced network there are two groups; moreover, ingroup members should share mutual friendships with each other and should have a mutual dislike of the outgroup. In the previous triad example (1c), p and q formed such an ingroup where they mutually disliked o. A generalization of this theorem to any number of ingroups will be discussed later in this paper. For now, please see an example and a nonexample of balance in a more complicated social network.



(a) Balanced Graph G (b) Unbalanced Graph U

Figure 2: A larger social network

Note that graph G can be partitioned into two groups, namely Group 1 $= \{A, B, C, D\}$ and Group 2 $= \{E, F\}$. This graph is balanced due to the fact that each ingroup member shares only positive edges and there are exclusively negative edges between groups, which satisfies the fourth statement provided in the Structure Theorem. Roberts provides an algorithm which assists in building these groups, although I will not include this in this paper.

Additionally, since G is balanced, it must be bipartite with only negative edges connecting between sets. Below is the bipartite representation of G, and the invalid bipartite form of U, where there exists a positive edge $\{B, F\}$ crossing from Group 1 to Group 2.

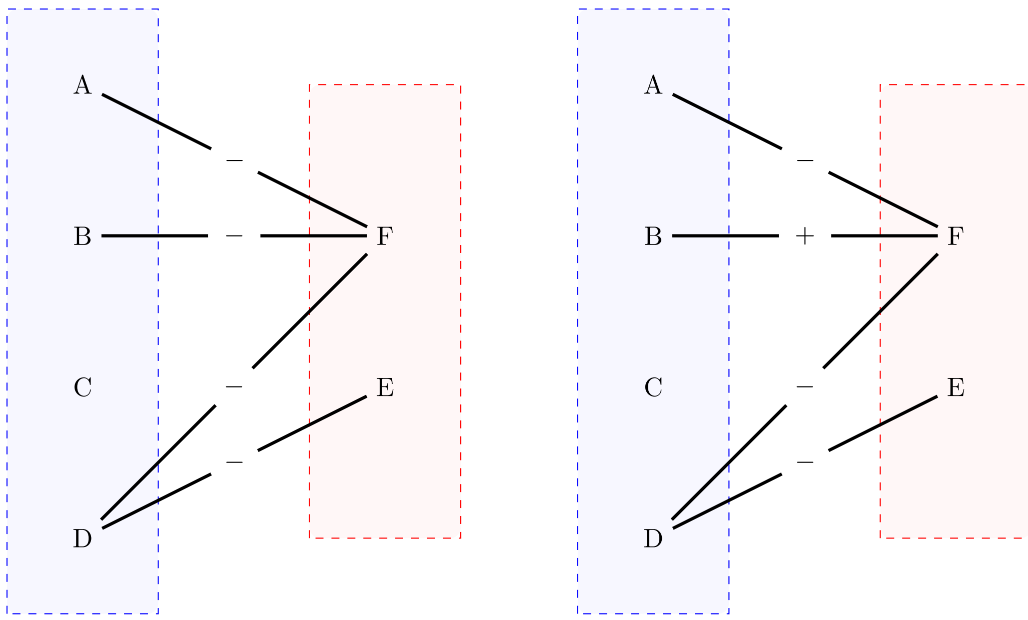
Intuitively, in graph G, those who are in Group 1 like each other, and those members of Group 1 who know members of Group 2 universally dislike them. In graph U, however, B shares a positive edge with F, which creates a conflict of interest between B and A, since A shares a negative edge with F.

One can also verify that graph G is balanced with the second and third statements from the Structure Theorem. To rigorously deal verify that G is unbalanced with statement 2, we must first introduce the notion of a spanning tree and the cycle basis.

Definition 1: A spanning tree in a graph G is a subgraph of G that includes every vertex of G and is also a tree. [6]

Statement 2 of the Structure Theorem requires that we verify that every cycle in G is positive; however, this may be unrealistic for large graphs. Instead, the spanning tree of G will help to identify some basis for which any cycle in G can be found through a linear combination of basis elements. This is vastly more efficient for verifying balance than enumerating every cycle in G—it's enough to simply check for a positive product of edges in the elements of the basis.

Definition 2: The cycle space $C(G)$ of a connected graph G is the collection of its Eulerian spanning subgraphs. [6]



(a) Balanced Graph G (b) Unbalanced Graph U

Figure 3: Bipartite forms of G and U

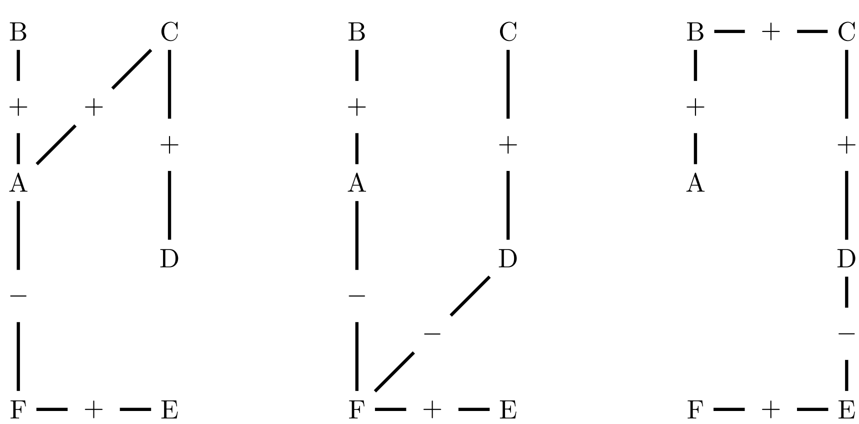
Recall that an Eulerian graph is a graph where one can traverse the entire graph passing by each vertex once and ending back at the start. Interestingly, since every cyclic graph has a cycle basis, one can decompose every connected graph into smaller Eulerian subgraphs. In order to make use of this basis, we can define a binary operation which constructs any cycle in G from two cycle basis elements C_i and C_j.

Definition 3: Given C_i, $C_j \in C(G)$, the symmetric difference is defined to be $C_i \oplus C_j = (C_i \setminus C_j) \cup (C_j \setminus C_i)$.

The symmetric difference essentially combines only the unique edges from each basis element. If they both share an edge, that edge is deleted. The symmetric difference the set-theoretic analog of the exclusive or (XOR) from Boolean arithmetic.

Theorem 2 (Cycle Basis) [6]**:**
Let G be a graph, and let $C(G)$ be its cycle space. Let T be a spanning tree of G, and for each edge e not in T, let C_e be the unique cycle in G that contains e and has all other edges in T. Then $B = \{C_e \,|\, e \notin T\}$ is a basis for $C(G)$.

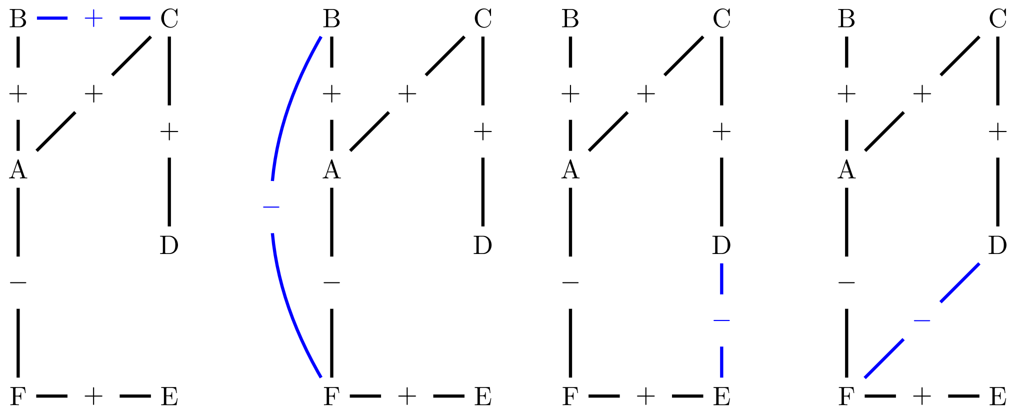
Finally, we can put these new items to use. First, we need to identify a spanning tree T for G. It's notable that many spanning trees may exist for a connected graph. Here are three such examples of spanning trees for graph G.



(a) Example 1 (b) Example 2 (c) Example 3

Figure 4: Example spanning trees

Then, we need to find all edges e from G which are not in T. Once a spanning tree is chosen, take note of the edges of the spanning tree T and the original edges of the graph G. To find the minimal set of cycles, one can simply connect the remaining edges from graph G inside T. With the first example spanning tree, note that we have the edge set $E_T = \{\{A, B\}, \{A, C\}, \{A, F\}, \{C, D\}, \{F, E\}\}$. Now, we must connect the remaining edges $e = E_G \setminus E_T = \{\{B, C\}, \{B, F\}, \{D, E\}, \{D, F\}\}$ to T, which are marked in blue below. Doing so yields the following four cycles which constitute the cycle basis of graph G.



(a) $T \cup \{B, C\}$ (b) $T \cup \{B, F\}$ (c) $T \cup \{D, E\}$ (d) $T \cup \{D, F\}$

Figure 5: Attaching missing edges e to T

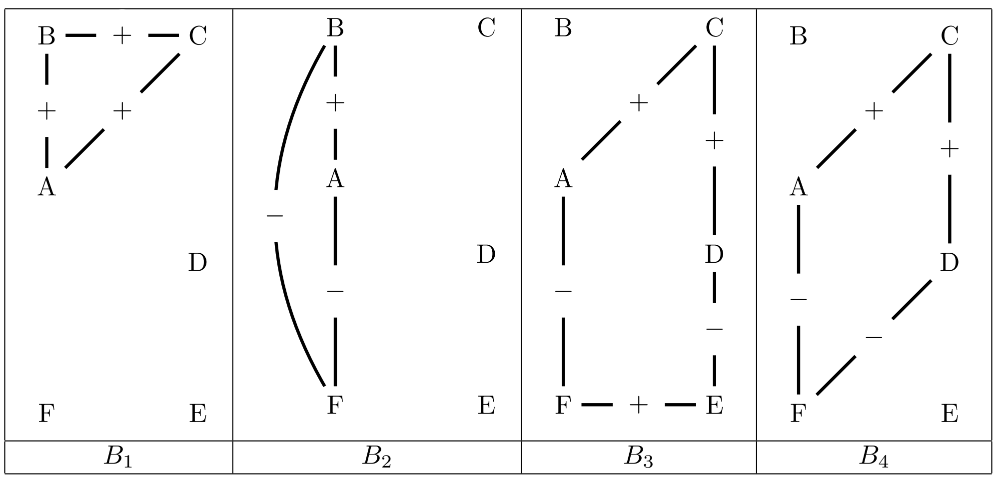
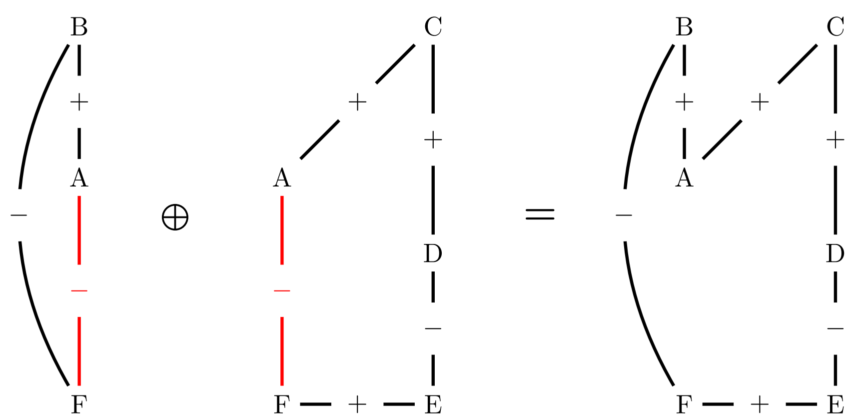


Table 1: Cycle Basis B of $C(G)$

Now that we have the cycle basis defined, we can construct any cycle from G by combining basis elements with the symmetric difference. For example, if we wanted to construct the cycle A, B, F, E, D, C, A from basis elements, we could compute $B_2 \oplus B_3$, shown visually below. The red edges are shared between the basis elements and are deleted in the final product.



(a) B_2 (b) B_3 (c) Desired cycle

Figure 6: Computing the symmetric difference

Interestingly, since the symmetric difference operates identically to the XOR operation, the cycle space actually forms a vector space over $\mathbb{Z}_2$ [3]. Here, the addition of two edges produces the number 2, but since the XOR operation is defined over $\mathbb{Z}_2$, that 2 "rolls over" into a 0, deleting the edge.

The cycle basis is useful for verifying that every cycle is positive. It's enough to show that if the signs of the basis elements are all positive, then every cycle in the original graph G is also positive. To see why, consider the following cases of computing the symmetric difference between two arbitrary balanced cycles C_i and C_j, and observe the behavior of negative edges. Since both cycles in the example are balanced, they both must have an even number of negative edges.

1. There are equivalent negative edges in C_i and C_j. The symmetric difference deletes them and produces an overall positive sign.
2. There are an even number of negative edges in both C_i and C_j, with C_i having $\geq 2n$ more negative edges than C_j, with $n \geq 1$. Here, any shared negative edges are deleted:
   (a) If an even number of edges are deleted, then an even number of unique negative edges will contribute to the resulting cycle, preserving the positive product.
   (b) If an odd number of negative edges are deleted, that leaves an odd number of unique edges in both C_i and C_j, therefore contributing an even number of negative edges to the resulting cycle, preserving the positive product.

In sum, since every $B_i \in B$ with $i \in \{1, 2, 3, 4\}$ is a positive cycle, we can safely conclude that all cycles in G are positive. Therefore, statement 2 of the Structure Theorem holds and G is balanced.

Addressing Weaknesses and Content Review

Roberts points out that there are several flaws with the simple model of balance: the model has assumes relationships are symmetric and lack a measure of strength, and the notion of balance is assumed to be "black or white," lacking any gradation [7]. Moreover, the model of balance necessitates that networks be composed of only two groups. What about when there are three or more groups?

To solve the symmetry issue, one could apply the concept of signed edges to digraphs, applying positive and negative signs to the arcs instead. This procedure creates an analog for the structure theorem with digraphs. To solve the strength issue, one could add scalar weights to the edges. One could define a weight $w \in \mathbb{R}$ for each edge which represents the strength of the association between two nodes; a weight $w \ll 0$ is extremely antagonistic and a weight $w \gg 0$ is extremely friendly. Additionally, the idea of weak balance allows for the graph to be partitioned into any number of groups rather than just two.

Since real life situations are unlikely to be perfectly balanced, perhaps the most useful tools are measures of partial balance. For this reason, I argue that Roberts should have included these partial measures as a key concept. The absence of modern methods could be explained by the publish year of *Discrete mathematical models, with applications to social, biological, and environmental problems*: 1976. Since then, researchers have progressively improved signed graph models to more accurately model real life. A positive comment I have on *DMM* is the format in which the content is introduced. Roberts first introduces the motivation of signed graphs from the social sciences before introducing the rigorous mathematics, and this influenced the flow of my paper heavily. The introduction of Heider's triads first, and the structure theorem later, eases the reader into the concepts with a steady pace. Moreover, Roberts provides myriad homework applications which are engaging for the student and help to illuminate the deeper ideas of the chapter. For example, one homework problem asks the student to map out the diplomatic relations in the Middle East and suggest which edges would optimally balance the network. Considering Syrian rebels just recently ousted Bashar al-Assad, which shifted the power distribution around Syria and weakened Russia and Iran, this practice problem is both engaging and relevant.

Since the chapter on signed graphs has a variety exercises—about half of the pages are filled with practice problems—I was inspired to research the applications more in depth. While I did not complete each exercise, each one gave an insight which aided in the research process.

There is some notable content I omitted from this paper. Primarily, I fail to describe in-depth the statements 3 and 4 from the structure theorem. While these statements stand at the core of the subject matter, each topic could have taken a page or more, and so I opted to focus heavily on statement 2 and examples. Moreover, I omit the a detailed discussion of n-partite weak balance and the structure theorem for digraphs. Had this paper been a longer project, I would have included these subjects in an extended section.

Finally, I was surprised by the degree of accuracy signed graphs provide in measuring real-life social phenomena. As will be discussed in the next section, signed graphs are incredibly useful in producing high-quality data of political coalition building. Researching for this paper brought forward impressive work from computational social scientists. The two most notable papers I found discuss coalition building in Congress and diplomatic power allocation in International Relations.

Further Applications

Signed networks can be applied to many situations in social science, including political science and international relations. In political science, increasing partisan polarization in the US is a well-studied phenomenon. For the purpose of modeling, balance in large political networks can serve as a proxy for polarization. Aref and Neal employed structural balance theory to analyze signed networks of legislators[1]. To construct the graph, the authors collected data of all co-authored bills in both the Senate and the House, and compared these to a stochastic model which served as the null. When two legislators shared less co-authored bills than the null model, they were connected via a negative edge. When two legislators had more co-authored bills than the null, they shared a positive edge. Now, the total balance of the graph will inform the researcher the degree to which Congress is polarized.

Since real-life social networks are unlikely to be balanced, the authors use two modified measurements which describe partial balance: triangle and normalized frustration indices. The triangle index measures the fraction of positive triangles (3-cycles) in the graph. The frustration index, on the other hand, measures the number of party-crossing co-authored bills—the same type of edge which violates the exclusively negative bipartition. Both indices therefore serve as a partial measure of balance. Finally, the authors corroborated prior studies which found that partisan polarization is increasing in the US [1].

Secondly, in international relations, researchers have modeled complex diplomatic relationships by applying game theory to signed graphs[5]. Countries are constrained both by their resources and their relationships when they try to maintain their geopolitical position. In the paper *Games on Signed Graphs*, Li and Morse construct a power allocation game between n countries within the environment of a unsigned graph. Firstly, they define the primitives:

1. The country index $\mathbf{n} = \{1, 2, ..., n\}$
2. An environment $\mathbb{E}_E = \{\mathcal{V}, \mathcal{E}_E\}$
3. $\mathcal{F}_i = \{\text{friends of country i}\}$
4. $\mathcal{A}_i = \{\text{adversaries of country i}\}$
5. $p_i \in \mathbb{R}^+ =$ power of country i, measured as some-present value currency
6. An alliance $\mathcal{S}$ with members $j \in \mathcal{S} \subset \mathbf{n}$ with $j \in \mathcal{F}_i$
7. Enemies of the alliance $\mathcal{A}_\mathcal{S} = \bigcup_{i \in \mathcal{S}} \mathcal{A}_i$

Each country has two disjoint sets which demarcate their friends and their adversaries. In the power allocation game, countries can choose to divide up their monetary power p_i to their friends and to the destruction of their enemies. They do so via a $(1 \times n)$ strategy vector u_{ij} subject to $u_{i1} + ... + u_{in} = p_i$. Allocations then are represented with a directed graph $\mathbb{E}_A = \{\mathcal{V}, \mathcal{E}_A\}$ which has the weights u_{ij}; negative signs are attached to enemies and positive signs are attached to friends. The game theoretic processes of the allocation game are beyond the scope of this paper. However, one key finding is that a country i can survive if it can satisfy either of the following constraints:

1. $p_i \geq \sum_{j \in A_i} p_j$: Country i has at least the same power as the total of its adversaries' power
2. $\sum_{i \in \mathcal{S}} p_i \geq \sum_{j \in \mathcal{S}_A} p_j$: Country i is part of a coalition $\mathcal{S}$ which has at least the same total power as the total all of the adversaries of the coalition's power

The takeaway from this model is that signed graphs provide the perfect tool for describing the ever-changing environment of international politics. The authors put it simply: countries pursue "survival and success in a constant or a changing environment, and may bring about some of the changes to the environment itself." For this reason, game theory on signed graphs is an ongoing field of research with great value to defense organizations worldwide.

References

[1] Samin Aref and Zachary Neal, *Detecting coalitions by optimally partitioning signed networks of political collaboration*, Scientific reports **10** (2020), no. 1, 1506.

[2] Abdelhakim El Maftouhi, Ararat Harutyunyan, and Yannis Manoussakis, *Weak balance in random signed graphs*, Internet Mathematics **11** (2015), no. 2, 143–154.

[3] Frank Harary, *Graph theory (on demand printing of 02787)*, CRC Press, 2018.

[4] Fritz Heider, *Attitudes and cognitive organization*, The Journal of psychology **21** (1946), no. 1, 107–112.

[5] Yuke Li and A Stephen Morse, *Games on signed graphs*, Automatica **140** (2022), 110243.

[6] Mary Radcliffe, *Cycle bases*, Carnegie Mellon, Department of Mathematical Sciences, 2018.

[7] Fred S. Roberts, *Discrete mathematical models, with applications to social, biological, and environmental problems*, (No Title) (1976).