INSTITUT EUROPEEN DES HAUT ETUDES INTERNATIONALES

MASTER THESIS

ALLIES PLAYING GAMES – PROSPECTS FOR A META-MODEL OF INTRA-ALLIANCE RELATIONS CASE STUDY: TRANSPACIFIC PARTNERSHIP STUMBLING OVER PYONGYANG'S BOMB

Supervisor
Prof. Dr Dagmar Roettches-Dubois

Bojan Savić
Date: 20-06-2008
Version: final draft

ISTANBUL-NICE-BERLIN-ISTANBUL, 2008

TABLE OF CONTENTS

1. INTRODUCTION

This paper will deal with a more 'exact approach' to international relations while trying to untie the complexity of politics essentially composed of two seemingly simple issues: war and peace. It will introduce a possibly insightful combination and intersection of two theoretical fields with quite distinct epistemological and ontological groundings: game theory and theories of alliances. This introduction will outline the research, defining (1) its object, (2) research questions relevant for the object of study, (3) research objectives, (4) relevant hypotheses, and (5) research methodology.

1.1 Defining the object of study: What is meant by intra-alliance relations?

The broad object of study is the *concept* and *reality* of international politico-security alliances. This term will be applied to all formal (based on treaties) and informal (without the legal base) modes of relationships between nation states who share common goals, and who cooperate with each other for a period of time in order to achieve those common goals. Their interests might overlap, but that is not necessarily a precondition for an international alliance (especially a short-term one), and while their goals can be jointly defined, they can also simply coincide. Contrary to a bulk of narrow definitions of international alliances (such as the one offered by George Liska in 1968), this paper does not claim that the indispensable element of any alliance is the mutual assistance pledge, as empirical cases prove the opposite, especially if informal arrangements (alignments[1]) are regarded. What is crucial, however, and what delimits the focus of this research, is an intended, purposeful and explicit cooperation among certain actors meant to bring about attainment of a common goal, delivering, in principle, individual benefits to all the parties involved. With respect to the scope of actors included, the framework and tools of this study are geared to tolerate alliances between states and specific non-state actors as well.

[1] This definition stems from the overall intention of the study to grasp the widest viable range of international agreements with politico-security agendas created to advance interests of their members. Therefore, whether an agreement is formal or informal is of a lesser importance. What matters is to determine the scope of *possible* components and determinants that drive those agreements.

More precisely, this applies to all armed groups, guerilla and terrorist formations whose operations directly affect regional and/or international security.

Even though the security role of international alliances has the crucial weight for this study, it is hardly conceivable that any alliance with a security agenda might not have more general political goals. As these overarching political goals wrap up the essence of the motivation and interests of allies, the syntagma 'international politico-security alliance' is adopted as the denotation of the object of this research in a broader sense of the term.

Furthermore, following the classical agency-structure problem, we can analytically dichotomize the ontology of international alliances and regard them either as operating environments of particular actors (*structures*) or as actors of world politics *per se*, with their internally defined goals and behavior (*agents)*. Although recognizing the limitations of this dichotomy, this study intends to dwell on alliances as *structures* hosting particular players, with individual interests, preferences, goals and behavior. This renders internal*, intra-alliance relations and politics the nucleus of the research, emphasizing intra-alliance bargaining and negotiations* on various issues, ranging from a day-to-day functioning of alliances to their long-term security strategies (if there are any), or *ad hoc* arrangements if the alliance in question is rather a short-term structure.

Another major feature of the object of study is the underlying approach taken to structure and dissect intra-alliance relations. This approach includes a comprehensive deployment of the apparatus of game theory, regarding members of alliances as players participating in cooperative and non-cooperative games. A consequential use of this method means qualifying allies' interests, motives for certain behavior, actions available and their goals as inputs to game strategies and strategic decisions. Therefore, the role of game theory will not be limited to its methodological functions, but will go much further. Its ontology and logic is to be the fundamental element of the study, taking its shape and reach beyond the standard analysis of applied game theory. This element is introduced as a whole theoretical input that will influence the purpose and objectives of the study, as formulating *the most effective way of studying intra-alliance relations through the application of game theory* is sought.

What follows from the previous definition is a twofold research: it concentrates on relations among members of international politico-security alliances; however, aiming to reveal how these relations, grasped as a theoretical concept as well, can be best built into the framework of game theory. Hence, intra-alliance relations are simultaneously regarded as (1) a class of empirical phenomena whose internal dynamics and features are examined (*the empirical level of the object*), and (2) as a theoretical concept discussed by all major IR schools of thought, and shaped through the use of game theory (*the meta-theoretical level of the object*).

It is crucial to emphasize that the meta-theoretical strand of this inquiry is to result in the construction of a unique *meta-model of intra-alliance relations* composed of two theoretical fields: theories of alliances and game theory. The goal is to design a meta-model abstract enough to embrace the whole class of phenomena defined as 'intra-alliance relations', and to provide the meta-model with tools that would help researchers and practitioners understand any given example of intra-alliance relations by primarily identifying components and determinants of those relations.

On the other hand, the empirical level of the object is to be completed through an explanatory and projection-driven study of relations within the alliance involving the United States, Republic of Korea (South Korea) and Japan. Special attention will be paid to the issue driving their relations, i.e. to the nuclear program of the Democratic People's Republic of Korea (North Korea) and its impact on relations within the Transpacific alliance.

1.2 Research questions

The study will be geared to answer two major sets of mutually dependent questions, each set corresponding to one of the levels of the object of study.

Firstly, the research discusses whether a meta-theoretical combination (a meta-model) of theories of alliances and game theory is possible, i.e. whether game theory can be applied to the study of intra-alliance relations. Secondly, it tries to discover whether that meta-model can (1) encompass detailed *descriptions* of the state and development of intra-alliance relations of any chosen international politico-security alliance, (2) *explain*

causes and effects of particular events or series of events, (3) *predict* future developments of any intra-alliance relations considered, and (4) serve as a tool for *policy prescriptions*. Thirdly, the research should answer if allies form their preferences on intra-alliance issues individually and independently, predominantly following their particularistic interests and preferences, or at the collective level that entails preference accommodation.

Concerning the empirical level of the object of study, it will firstly seek to reveal the way relations among the US, South Korea and Japan influenced their decisions on how to deal with the challenge of the North Korean nuclear program in two crises induced by North Korea's decision to withdraw from the Nuclear Nonproliferation Treaty (1993-1994 and 2002-2007). Secondly, it will offer a cautious prediction on the future preference of the Transpacific Alliance on the posture that should be taken towards North Korea. This future preference will, if the logic of this study proves valid, depend both on North Korea's behavior and relations within the Alliance.

Noteworthy is that addressing both portions of research questions is mutually dependent. Answering the meta-theoretical set of questions results in creating a meta-model that is applied to the case of relations within the Transpacific Alliance. Pondering the response of the US, South Korea and Japan to the North Korean nuclear challenge is, in turn, facilitated and structured along the lines of the meta-model, testing its applicability and weight.

1.3 Research hypotheses

This research will operate under a set of hypotheses of different levels of generality. These hypotheses are strongly related to the research questions and objectives stated above, they are dependent upon them, and are produced following them logically. Hypotheses are divided into a set related to the meta-theoretical strand of the research and the ones resting on its empirical portion.

Firstly, the most general proposition of the *meta-theoretical* strand of this research is that game theory's tools are highly applicable to the study of intra-alliance relations, as its five defining elements (players, strategies, information distribution, payoff consequences and preferences over payoffs) are entirely applicable to the inquiry of

components and determinants of international politico-security alliances ($_{G}MTH_{1}$)[2]. Secondly, allies tend to formulate their preferences individually factoring in an independently perceived collective (alliance-level) interest ($_{G}MTH_{2}$)[3]. Thirdly, international politico-security alliances exert an independent influence on the formation of those preferences, yet predominantly at the individual level of alliance-members ($_{G}MTH_{3}$)[4].

These propositions, in turn, rest on two specific hypotheses ($_{S}MTH_{1}$ and $_{S}MTH_{2}$)[5]. Firstly, alliance-level preferences tend to be formed following the intra-alliance preference accommodation procedure, mathematically defined as an alliance-conditioned utility function, rather than by the preference aggregation technique ($_{S}MTH_{1}$), which is a hypothesis in accordance with $_{G}MTH_{3}$. Secondly, an alliance-level preference cannot be simply disaggregated into its individual preference-components ($_{S}MTH_{2}$).

As for the *empirical* strand of the research, the most general hypothesis is that the Transpacific Alliance will try to maintain an overtly tough and firm posture towards North Korea accompanied by a harsh public discourse, whereas in practice it will continue with the policy of intense efforts to involve and particularly commit North Korea to talks with International Atomic Energy Agency and South Korea ($_{G}EH_{1}$)[6]. Secondly, this transition is an alliance-level preference reached through a preference accommodation process ($_{G}EH_{2}$)[7].

1.4. Methodology

A set of methods that allow for a collection, classification and processing of data are needed to answer questions enlisted above.

---

[2] General meta-theoretical hypothesis no. 1
[3] General meta-theoretical hypothesis no. 2
[4] General meta-theoretical hypothesis no. 3
[5] Specific meta-theoretical hypotheses no. 1 and 2
[6] General empirical hypotheses no. 1
[7] General empirical hypothesis no. 2

The essential method to be deployed is '*meta-modeling'*, i.e. a deconstruction and reconstruction of existing theories to form a new theoretical framework capable of most suitably explaining any accordingly chosen phenomenon. This method will mutually adapt the theoretical frameworks of game theory and theories of alliances into a single research apparatus. As a sort of abstraction of already abstract models (therefore the name of the method), meta-models are mostly used in computer science, software and systems engineering (known as metadata modeling), and neuro-linguistic programming. However, in this research, the meta-modeling method is adapted to the quality and scope of Political Science theories, which makes this inquiry a methodological breakthrough, too. Its deployment will practically mean:

- Presenting theories relevant for the scope of this study (i.e. theories of alliances, on the one hand, and discussions on game theory, on the other hand);
- Deconstructing the substance of the theories offered into their components (hypotheses and arguments in their favor);
- Abstracting crucial features of the theories presented following the criterion of scientific relevance;
- Separately framing two inputs of the meta-model by congruently merging the elements abstracted into a logically and substantially coherent set of analytical instruments ('input1' drawn from theories of alliances and 'input2' from game theory) ;
- Identifying components and determinants of intra-alliance relations, based on results of the examination phases depicted above;
- Mutually adapting input1 and input2, following the criterion of logical and substantial coherence, and thus completing the meta-model.

Moreover, such a fundamental analytical method as *abstraction* will serve to emphasize crucial and for the purpose of this study most relevant aspects of particular theories while framing the inputs of the meta-model. Some other fundamental methods (e.g. *synthesis*, *concretization* and *generalization*) will help reach the general objectives as previously stated.

Operative qualitative methods (*case study*) will be used to test whether meta-modeled theories prove applicable to empirical situations. These methods are also meant to help identify potential limitations of the meta-model.

Comparative analysis will add to the application of the meta-model by contrasting features of relations within the Transpacific Alliance with the components and determinants of intra-alliance relations as abstractly defined in the meta-model. Furthermore, comparative analysis will serve to detect deviations between the calculated (predicted) and actual outcomes of all analyzed games that serve as examples supporting the construction of the meta-model.

Techniques specifically belonging to the methodological framework of game theory will be used as well, such as '*tactical analysis' methods* i.e. *evaluation function* and *search technique*, additionally rearranged by pruning technique, and *genetic analysis* in order to (re)construct policy-making patterns, strategies and bargaining interplay of players. *Backward induction* will be deployed to reveal nuances of players' reasoning and to explain their subsequent actions. Most importantly, game theoretical *functions* (e.g. alliance-conditioned utility function) and *procedures of preference formation* are meant to implement abstract averments of the meta-model. For instance, game theory's preference accommodation procedure puts into practice the abstract proposition of the meta-model that alliance-decisions tend to be taken through the policy accommodation process of intra-alliance bargaining.

Underlying the aforementioned tools, relevant *literature search* will be used as a method of data collection.

2. THEORIES OF ALLIANCES: FRAMING INPUT1

The goal of this section will be to provide readers with a set of tools that will prospectively construct a final model of the analysis of intra-alliance relations.

The purpose of this section will thus be to practically operationalize the usage of the phrase "intra-alliance relations", vastly used in the study.

Before starting with an in-dept examination, two observations are to be made at this point. Firstly, framing the substance of this portion of analytical input will pursue an examination of crucial theoretical concepts rather than just a mere overview and description, thus consciously running a risk of disregarding some of the less relevant[8] aspects of the theories in question. However, this dissecting approach will give a clearer image of the matter dealt with apart from meaningfully rationalizing the extent of the argumentation. Another remark is that the range of concepts that will be examined and possibly adopted in the course of the study will come from a variety of theoretical backgrounds, all being based upon different presumptions about reality and its constituting features, which, consequently leads to distinct explanations, projections and policy recommendations. This wide range of normative stances will however be moderately abstracted (as a complete reduction of certain factors can only lead to a stupefaction of theories in question, rendering them utterly useless), while trying to adopt rather neutral and interdisciplinary communicating properties of the concepts observed.

2.1 Which theories of alliances – overarching notions

2.1.1 Kenneth Waltz – seeking security under anarchy

A major challenge to our "value-neutrality" and objectivity-inclined approach (Weber in Gerth & Mills, 1946) is posed in the very beginning by the question: how to cope with such strongly pronounced methodological contradictions in the study of world politics as allusively displayed in the debate between classical (traditionalist) and "neo" theories of IR: where should the study start from? Where do the causes of world politics

[8] In the context of this study

rest? Could it be in the system and its structure or is the state (or a lower level)-unit that predetermines its tectonics? This dilemma will persistently play a role while an attempt of dissection of Kenneth Waltz's observations on international alliances is made.

In accordance with Waltz's general viewpoints on the nature of international relations (i.e. his systemic and neorealist approach), a few ideas specifically applicable to our case are offered, and they all carry in the essence of Waltz's stance. Firstly, we should keep in mind that he sees international alliances as a mere convergence of two or more states' short-term interests, and, secondly, that the purpose of alliances is exhausted in a state's individual security aggrandizement. The short-term destiny of alliances and their dissolvableness apparently spring out, according to Waltz, from the ever-applicability of the self-help principle (Waltz, 1979, p. 126). Additional points related to these two main portions of discussion will be added.

However, the first half of Waltz's argument introduces a set of edifying concepts in terms of this study. The core of the argument revolves around the notion of interests in international affairs, most generically defined as 'survival'. A situation in which such two or more compatible interests for survival can reach a quasi-institutional or full-fledged institutional rapprochement, fostered by a prospect of a common challenger (or a common threat), leads eventually to a causal chain that straightforwardly depicts the internal logic of alliances: converging interests in a strictly defined issue-area leading to a recognition of states' interdependence, in turn leading to a process of cooperation. However, this interdependence is and, rationalistically regarded, must be subject to revisions as soon as the structure of the system suggests so by practically compelling states to behave uncooperatively. Therefore, Waltz's notion of cooperation and interdependence has a substantially different meaning then the one assigned by Keohane and Nye (2000), as discussed later. Inferring from his postulates, a limited extent of interdependence is a consequence of converging interests, not an independent cause or a push for cooperation. There is no such interdependence that could outshine the imperative of survival and make a state cooperate if it is contradictory to its core interests, i.e. if the benefits of cooperation are outweighed by its costs[9]. Therefore, the hypothesis on alliances framed in this way encompasses the concept of interdependence

[9] We will refer back to the problem of a procedural definition of rationality in the third section of this study

only as an intervening, intermediate variable that helps us explain the link between interest (independent variable) and cooperation (dependent variable). It does not possess any genuine or autonomous explanatory value *per se*. That in turn means that alliances have no internal properties that can compel states to cooperate. They cannot draw their *raison d'être* from structures interactively defined between (or among) states based on their enduring internal characteristics. Here, we see that an externally (systemically) determined entity of national interest paves the path of individual state's behavior in the international system, leaving quite a limited space for a genuine influence to such entities as international regimes, and among them, international alliances. Alliance is an effect of the system, not its determinant. There is no such feature or internal mechanism (supposedly leading to alliances, such as interdependence) beyond the realization of unit's interest that can lead to its formation. Alliances exist within the structure, they are a part of it, but they are not its constitutive element. Therefore, they can be regarded as a decomposable condition of capabilities distribution, which is a true element of the structure's definition (Waltz, 1979, p. 100). This argument was sketched concisely by Krasner in his famous diagram, herein reproduced as Figure 1.

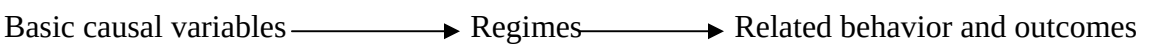

FIGURE 1. *International regimes as an intervening variable* (Krasner, 1983, 5)

Consequently, the notions that our study inherits from this first part of Waltz's discussion are:

- Alliances are built in a context of a predetermining anarchical structure of international system;
- That structure (a major feature of the system) entails the actual interests (as practical meanings of the overarching urge to survive) of units that may, under favorable circumstances, cooperate in an issue-area, thus forming an alliance;
- The alliance will last as long as it fulfills the purposes derived from the constellation of interests engaged.

As it attributes no independent causation capacity to concepts such as interdependence, cooperation, institutions and regimes, Waltz's study does not offer a scientific rationale

for their inclusion as inputs of our meta-theoretical method. Nonetheless, the course of the study will show the importance of these notions following some differing logics of world politics later on.

The second part of Waltz's argumentation sheds some light on the meaning and purpose of alliances by introducing three specific methods of security attainment. For the purposes of this study, we will only emphasize that building alliances in an anarchical international system is a method of increasing participating states' respective individual securities so they can more successfully balance other competing states in the system (either by defensive or offensive undertakings) perceived as threats. This method of balancing achieved through an external capabilities enlargement has its internal counterpart as well, i.e. an autonomous, domestic increase of certain state's overall (primarily military) capabilities[10]. The final, third element of conquests/conflict as a method of power maximization is added to the previous two, revealing that states can aggrandize their security through a maximization of power by capturing territory and resources of other states or by attaining concessions in the post-conflict bargaining environment. However, all the three are methods needed (and at disposal) for an effective counterbalancing to be executed. Noteworthy is that the first external method directly leads to alliance formation, as other actor(s) in international system are invited to join a foreign policy coordinating arrangement (which an alliance definitely is according to Waltz). If we look at the essence of this argument, it fundamentally asserts that while seeking security, states form military alliances to upgrade the total sum of power on their side of the equation[11] needed to balance contenders in the system. Although these methods of security maximization are complementary (and can happen simultaneously), the internally driven one is preferable, as it does not limit particular state's autonomy, usually bounded in an international arrangement, and brings less risks then waging wars . However, it is generally not sufficient to render the state in question internationally autarchic, which is why it has to form alliances (and/or engage in a conflict), primarily in

[10] We will later see how this factor influences alliance formation and maintenance.
[11] Although not their sole individual power, as it gets enlarged only by following the other path of internal capabilities augmentation.

order to *shift the capability distribution* across the system[12] (Waltz, 1979), i.e. to add an amount of power to its own potency. Due to the nature of this method of security attainment, this model of international alliances construction is also interpreted by the scholar community as the capability aggregation model (Morrow, 1991; Kimball, 2006), thus correctly describing its nature and purpose.

What learnt from this portion of Waltz's argumentation is that the notions of *security* and *power are in the ontological core of international politico-security alliances*. Even if we step out from the (neo)realist assumptions on international politics, we can hardly deny this major source of motivation for states to enter international arrangements[13], although we could with comparable difficulty claim it would be the only one (Kimball, 2006), at least if we understand security in exclusively military and hard power terms.

However, as the issues of power and security are deeply interwoven in the constitution of alliances, we will adopt them as concepts prospectively constructive for our grand model of intra-alliance behavior based on game theory. Dimensions of power and security remain identifiable in a wide range of international issues, be it economy, ecology or a constellation of military threats. For that reason (which fulfills the theoretical condition of generality), we decide to enrich our model by including these two concepts, although the exact modality of that inclusion remains a task of a latter phase of our meta-modeling method. By doing so we reasonably abstract from the imperative and a rather narrow meaning of power and security, defined by realist and neorealist schools as 'hard security' and 'hard power'. Given our intention to overcome the limitations of particular theoretical schools, the flexibility of the meta-model we are building, and the compatibility of new ontology in the study of IR with this meta-model, we accept broader definitions of security and power as both hard and soft[14].

[12] Nevertheless, this alliance-contributed shift is not structural, but momentous and issue-limited, as claimed in the first part of Waltz's argument paraphrased here.

[13] Moreover, if those arrangements are security-driven, then this circumstance becomes quite tautological and pointless to contradict.

[14] For a definition, see Buzan, B., Weaver, O. & De Wilde, J. (1998). *Security: A New Framework for Analysis*. London: Lynne Rienner Publisher

Building upon the previous argumentation, we can now discuss how the notions of individual interest, power and security aggrandizement entail the life cycle of alliances. This aspect of analysis attaches a dynamic facet to the previously defined stable structure. Conceptually, the problem is posed at the level of alliance management. "In quest for security, alliances may have to be made. "Once made, they have to be managed" (Waltz, 1979, p. 166). The formula of this alliance management, according to Waltz, lies in a dynamic interest intersection. This idea, as developed by the author, is a natural extension of the two previous concepts-based arguments on (1) the constitutive role of interests, and (2) power-security nexus. Furthermore, its grounding is consistent with the overall theoretical context: managing alliances is nothing but conceiving your own state's logical interests inherent to a directly capability distribution-driven system. As the system is competitive, besting your opponents is done through a pragmatic process of interest identification. However, as the prime goal is to survive by deterring or compelling the enemy, the core of each alliance is composed of an ordinarily negative interest: fear of other states. At this point, it also becomes clear that the role assigned to alliances by Waltz is predominantly defensive, which is not a major surprise, given a consensual classification of Waltz as a "defensive neorealist" by the scholars' community in general (Taliaferro, 2000, p. 135). In accordance with the aforesaid, an alliance is facing a serious challenge whenever positive interests of allies become more pronounced than the fear that brought them together. Put it straightforwardly, what they want to achieve individually does not necessarily follow from their shared fear. Therefore, the reality of diverging interests is never absolved; it just might be temporarily hushed[15]. Although ostensibly well-founded, we find this view of alliances rather narrow. Even though Waltz leaves some limited room for a positive interests-driven alliance, we see this stance as not completely matching the reality: cases of e.g. offensive alliances[16] (if

[15] Historical examples are strikingly numerous: Athens and Sparta in the Greek-Persian Wars, Russia and UK on the eve of the World War I, European balance of power system throughout XVIII and XIX century, US-Taliban tight alignment in Afghanistan in the 1980s etc.

[16] Possible researches using statistical methods to establish the relative frequency of offensive alliances would have to take into account a definition problem posed at this point: offensive alliances with positive interests would have to be differentiated from offensive alliances created in preemptive purposes (i.e. with basically negative interests).

we simplify the issue for the sake of clarification) are not missing either[17]. A specific theoretical counterargument is, amongst others, offered by the so-called "extended power transition model" framed by Woosang Kim, where he consistently argues that alliances have a prominent offensive role in challenging the power of hegemon, or a dominant power in the system by encouraging the challengers (1991). We would just add: and not only then.

Despite quite a mechanical and mathematical model of alliances (where capabilities are being aggregated), Waltz recognizes structural realism (and systems approaches in general) cannot truly predict responses of each state to systemic stimuli. Therefore, a continuum of determinism is broken, but in a way that, contrary to widespread misperceptions of neorealism, leaves room for certain (even considerable) autonomy of states when it comes to specific actions. One of those actions is choosing actual allies with specific properties, features and assets. Another is choosing an intra-alliance strategy once the regime is formed. However, these strategies are entailed by systemic conditions, and cannot afford to ignore two simple facts embracing international politics: an uneven distribution of capabilities and the lack of central government. Waltz particularly focuses on three strategies:

4 Whether to endorse allies' foreign policy actions and join them even if they might look as adventures (and they are still within the issue-area defined in the founding agreement), just for the sake of alliance preservation (the strategy of chainganging);

5 Whether to let weaker allies decide on an equal footing with their stronger partners in alliances among unequals;

6 Whether and how to timely and optimally deploy the wide-known strategies of free-riding, buck-passing, chainganging and bandwagoning (Waltz, 1979).

As the dilemmas of specific strategies do not hold the sufficient relevance for the substance of this study, their broader and further detailed discussion will be left to more operationalizing and case-study oriented papers. Given that this paper aims at constructing a model of analysis and policy prediction, concrete strategic or tactical considerations would be redundant at this point, as they would not affect or

[17] The example of Axis powers is just one of them, and probably the most notable one (negative interest related to a common fear of communism is hardly a convincible argument).

constructively enrich the *framework* that is being built up. Theoretical modeling takes into account the context of certain concepts, their valence (combining power), employability and core facets, which is why examining the specific implications of the strategies defined remains futile in this sense. However, underlying that all these strategies are compatible with a consequential usage of game theory is crucial. Then, it is up to particular studies of international relations to examine the concrete dimensions of applicability of the model developed (as a brief case study will demonstrate in section no. 4).

As shown in the third part of the examination of Waltz's alliance-related views, the concepts we have discussed and prospectively intend to incorporate into our final model (if the condition of coherence is fulfilled) can be summarized as following:

- Alliance management, preservation and functioning depend on the *vector of interests* involved in an issue[18];
- Interests at stake influence allies' *internal strategies*, i.e. sets of actions available to allies at a given moment.

Therefore, we can emphasize once more that the notion of interest represents the focal point and ontological core of international alliances, determining them directly in the light of internationally experienced anarchy.

2.1.2 Expanding the rational choice-rooted approach

As we have shown, the essence of the neorealist arguments is that each state's relative gains calculations mercilessly dictate the life pace of any interstate arrangements, including politico-security alliances. However, the range of these arguments remains quite general, indeed corresponding with the 'grand level' of the theory in terms of IR. Other, middle-range theories are therefore needed to examine the study object in more details and offer more specific insights.

A rather safe approach is to simply ask: What are those concrete benefits that make states engage politico-militarily? Answering this question for the purposes of this

[18] Vector of interest being defined by its three features: course, magnitude and direction.

study presupposes actually identifying a couple of tangible gains that induce interstate joint security endeavors.

A seminal work with this regard is Altfeld's rational choice model of security alliances (1984), constructed to show how states gaining additional security by entering alliances have to sacrifice their absolute foreign[19] policy autonomy. By forming alliances, they show how at a specific moment, trading a portion of that autonomy is valued less than the security and power capability subsequently acquired. The group (alliance) outcome is increased security for everybody and everybody's decreased decision-making autonomy, due to a necessary policy coordination, bargaining process and mutual acceptance of joint actions. Therefore, the classes of allying benefits are supposedly equally or more correctly put, symmetrically distributed.

Departing from this straightforward model, James D. Morrow presents a more nuanced security-autonomy trade-off theory of international alliances to encompass both the cases of symmetric alliances (where all allies gain security) and asymmetric ones, where one ally (or one group of states within alliance) gains security and the other (or others) autonomy[20]. This theory regards defensive neorealism's capability aggregation motives to enter into alliances only as a special case, one of many, thus pledging to offer far broader and more detailed insights into alliance formation practices simultaneously (1991, p. 905). In terms of our inquiry, it is particularly perceptive to see how Morrow defines security and autonomy, as these concepts are widely used in our discussions.

He defines a nation's individual security as "its ability to maintain the current resolution of the issues that it wishes to preserve" (Morrow, 1991, p. 908). It is hard not to conclude that this definition of security is rather limited, as it also bears the label of the problematic defensive dimension, disregarding that certain states can *feel* (and indeed felt) secure throughout history[21] only if their capabilities are being constantly expended and the status quo revised. To say that the "resolution of issues" is more a tendency then an absolute temporal change is logically tense, especially as it leaves no room for an

[19] and sometimes even domestic

[20] As, for instance, the Warsaw Pact gave security to Eastern European communist regimes (used to equal the whole of the state) and gained autonomy out of that (provided Soviet's army access to the Cold War contenders in Western Europe, thus facilitating its power projection)

[21] This calls for a reference to the famous thesis of Paul Kennedy on the notion of imperial overstretch and its causes, developed in his influential book The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000 (1987).

initial revision of state of international relations. Therefore, the conservatism of this definition warns about its limited applicability.

However, when we look at his definition of autonomy, we manage to find 'the missing element' of the previous definition, as it is stating that "a nation's autonomy [is] the degree to which it pursues the desired changes in the status quo" (Morrow, 1991, pp. 908-909). Taking them jointly, we see that their separation is slightly artificial, especially if confronted with more comprehensive definitions of security[22]. Nevertheless, the rest of the article shows that the intended implications of these definitions are tightly connected with the meaning we have previously adopted, rendering this article still potentially relevant for our study.

Morrow further on maintains that national interests can be dichotomized into security and autonomy seeking. Therefore, given either security or autonomy constant, increasing the other is generally preferred, which means that states have convex (i.e. basically inversely proportional) preferences over autonomy and security. The ultimate interest of a state is to, considering this convex distribution of preferences, keep small overall difference between the satisfactions of the two. States ideally want to trade as little autonomy for security as possible and *vice versa* to reach a state of balance where they remain both secure and autonomous enough in their foreign policy. Of course, the exact assessment and evaluation of how much security/autonomy is needed ultimately depends on a specific country's definition of national interest. Changes in autonomy are reflected through a state's capacity to with more (or less) freedom of choice and with a wider (or narrower) scope of strategy options influences the state of international relations (talking about an increase of autonomy, it means using military bases of its allies, influencing their domestic decision-making, increasing its power projection capacities etc.), whereas changes in security designate an either increased defensive or offensive capacity of the state in question, or a decrease of security (if it is traded for larger autonomy), as its new position may attract new challenges and risks coming from other nations, and because the state in question will undertake actions to defend its allies and involve with conflicts it would otherwise stay out of. Finally, its security is decreased because it is now sharing the absolute value of its capabilities with other states who

[22] Such as Copenhagen School's definition of security in terms of identity (Williams, 2003)

offered their autonomy as an alliance asset, but not the military capabilities themselves. Examining specific cases, he notices that an asymmetry in trade-offs correlates, and is caused by the asymmetry in power capabilities (which makes alliances made of stronger/weaker partner groups of states), and vice versa. Testing this framework (and the hypotheses derived from it) on a set of international military alliances formed between 1815 and 1965, Morrow draws the following conclusions:

- Asymmetric alliances will be easier to form, and they tend to last longer than the symmetric ones;
- Regardless of the empirical type of alliance, the more significant shift of power within it, the more likely it is for the alliance to break;
- Second-rank major powers will be more inclined to form asymmetric alliances as their power increases (Morrow, 1991).

Incorporating Morrow's insights into our structure will mean drawing relevant and general conclusions, and making them coherent with the so far constitution of our meta-model. Hence, the input we frame at this point is:

- Gains drawn from entering into alliances do not have only the narrow security meaning, they encompass at least one more aspect of national interest - autonomy[23];
- Managing alliances presupposes a constant bargaining and adjustment of reciprocal benefits and costs, bargaining deadlocks entail an instability of alliances, as the trade-off, the engine of alliances is impeded;
- Alliance members should be regarded as dynamic actors whose individual characteristics change over time, thus affecting life of alliances;
- Alliance members' shifts in individual features might both foster and discourage alliances' development and survival, depending on particularities of each case;
- Generally, the greater discrepancy and asymmetry in interstate capabilities, the greater capacity for an alliance to exist – partnerships of equals are hardly achievable.

[23] Other theories add more benefits different in their nature, but not necessarily incompatible.

2.1.3 Enriching the mosaic of alliances: specific strategies, or how allies think

Given Waltz's own remarks (Waltz, 1979, p. 39) and general assessments of his theory (Taliaferro, 2000; Christensen & Snyder, 1990), it is a well-established opinion that structural realism can hardly explain and predict foreign policy actions and behaviors of specific states in concrete situations, mainly because systemic-structural constrains and determinants act only indirectly and contextually, although comprehensively. However, precisely this kind of undertaking is needed for a meaningful intersection of game theory and theories of alliances, not just an unfounded speculation. Because, if the final model is incapable of predicting (even tentatively) behavior of the actors involved, it cannot fulfill its twofold imperative teleology comprising explication and projection either, i.e. its practical applicability and scientific reach remain limited. Furthermore, as game theory focuses on the behavior of units (actors or 'players', if its parlance is used) as game 'inductors' and engines, failing to frame an input of theories of alliances that would do the same will mean a major deficiency of our model, rendering it incoherent and deprived of consistent analytical tools.

To compensate for this exigent lacuna, we will now more attentively focus on incorporating the discussion about four main intra-alliance strategies into our raising model. Explicitly, the strategies of *buck-passing, chainganging, bandwagoning* and *free-riding*[24] will be deconstructed into their logically composing elements, and then re-included into a new theoretical surrounding, so to become a natural and coherent extension of structural realism's views on intra-alliance relations and nature of interstate alliances.

Making a scheme of the link needed, we will first pose the question and dilemma before us: How to explain specific foreign policy decisions of alliance-members that stand in a stark contrast with expectations derived from structural realism? Also, we keep in mind that our elaboration should not only include an understanding of 'surprising' and

[24] For a detailed explanation of the strategies, see Christensen, T.J. & Snyder, J. (1990). Chain Gangs and Passed Bucks; Predicting Alliance Patterns in Multipolarity. *International Organization*, 44, 2; and Wallner, K. (2002). The Provision of Public Goods in International Relations: A Comment on "Goods, Games, and Institutions". *International Politics Science Review*, Vol. 23, No. 4, pp. 393-401

'unpredictable[25]' events, but should also be equipped to clarify all cases occurring under same (or highly similar) circumstances. The link between systemic approaches and unit-level examinations lies here in the factor of *perception*, making thus an inevitable link with Robert Jervis' work on security dilemma (Christensen & Snyder, 1990). Logical and substantial links with structural realism are preserved in the systemic conditions posed upon the actors (anarchy, self-help system, capabilities distribution, and competition). However, the notion of perception (and, conditionally said, misperception) should explain why alliance members do not always seek to aggrandize their power, i.e. why they risk even large-scale conflicts although the according level of systemic pressure lacks[26], or incentives for an offensive behavior are likewise missing.

Without getting into details of their fruitful discussion[27], we will just emphasize that Christensen and Snyder (1990) show through an examination of historical examples (situations before the two world wars with a range of actors involved) that the role of perception of threat posed by other states is crucial, just as in the explanation of security dilemma, i.e. states cannot often discern defensive from offensive capabilities and they can, moreover, be never sure about other states' intentions of how to use the military capacities possessed. With alliances, this leads to various forms of perverted behavior of allies. They either underestimate the threat individually and let their allies deal with it on their own (which at the end jeopardizes the members of the alliance, as it is seen as a single actor by the attacker), or uncritically follow the perception of threat as estimated by their allies, thus risking larger conflicts. The former strategy case of buck-passing and the latter of chain-ganging, combined with the concept of perception, display the negative effects of networked responsibilities and rights within alliances, which makes this explanatory structure vastly distinct from a mere application of security dilemma theory. Namely, one of the effects of shared and mutual commitments is a considerable room

[25] In terms of Waltz's defensive neorealism, i.e. structural realism

[26] Structural realism's, i.e. defensive neorealism's *systemic pressure-individual state's role- another state reaction* mechanism can, as we saw criticizing it, grasp only one aspect of offensive alliances. Beyond that framework, the dilemma of how to incorporate the logic of offensive alliances remains. Nevertheless, it remains so only seemingly. The notion of perception represents an alternative explanation to pure motives of resource expansion, whereas we clearly stated that we do acknowledge the influence of various levels of actors on world politics, not just the systemic level. This allows for a possibility of certain state inducing conflicts for non-defensive reasons.

[27] For more details, see: Christensen, T.J. & Snyder, J. (1990). Chain Gangs and Passed Bucks; Predicting Alliance Patterns in Multipolarity. *International Organization*, 44, 2

created for both responsibility and blame shifting, and uncritical conflict involvement for the sake of alliance preservation and obligations fulfillment. On the other hand, these ostensible shortcomings of alliance structure are a regular factor to be counted with. Here, we will digress from the course of argumentation of the authors to make a remark of our own.

Alliances widely experience the above temptations of buck-passing and chain-ganging; still, it is a matter of state's skillful foreign policy management how to make fruitful strategies out of these obstacles. If alliance members are tempted to buck-pass the problem over to their allies (deteriorating the significance of mutual commitments), or if they are occasionally brought into conflicts by *predominantly*[28] fulfilling their responsibilities or by overreacting and recklessly inheriting their allies' threat estimations[29] (overestimating and overvaluing[30] those commitments), they might even manipulate these temptations and use them as their own policy instruments (in order to save domestic resources, preserve a larger alliance potential etc. by using the strategy of buck-passing, or to increase the credibility of threat posed before third parties, promote themselves as respectable powers, promote their foreign policy goals etc. by using the strategy of chain-ganging). This, however, presupposes a carefully planned dosage of deployment of these strategies, so to avoid the alliance breakup or an ultimate failure before the common goals are accomplished.

Except for containing additionally pronounced economic factors in the cost-benefit analysis of decision-making, the other two strategies of bandwagoning and free-riding can in terms of contextualization be compared to the ones previously depicted (free-riding to buck-passing and bandwagoning to chainganging)[31].

The point drawn from this portion of analysis is that controlling and adjusting your own perceptions of systemic actors' capabilities, and meaningfully manipulating other states' perceptions of your foreign policy contains an immense potential for foreign policy engineering, as well as for prescriptions and projections. These are the tools alliance-members use to spread their control over their both friends and foes (supposedly

[28] motives are rarely and hardly singular
[29] Possibly directly inferable from their particularistic interests
[30] Also depends on the character of the conflict they are being involved in, i.e. whether it is peripheral or central to the system (Christensen & Snyder, 1990, p. 142)
[31] However, this analogy is not made by Christensen and Snyder.

outside of the alliance). Nevertheless, while constructing their foreign polices within alliances, states will only do so under the recognized systemic constraints.

The section has moderated the analytically perilous one-sidedness of Waltz's structural realism[32], enriching it with conceptual observations on the potential of units to determine the state of intra-alliance relations by deploying diverse strategies of foreign policy.

It has also shown that alliance-members do not have a singular agenda or a singular modus of behavior (resting on diverging interests) that is in accordance with the founding agreement, which means that vivid intra-alliance politics are permanently reoccurring. This observation in turn poses high demands on dialectical and dynamic dimensions of our behavioral intra-alliance model's explanatory power, requiring a tentative and comprehensive inclusion of processes and their cycles into it. It also announces a multiple fruitfulness of the game theory analytical apparatus in this meta-model, as it opens dilemmas of intra-alliance policy coordination, information exchange and cooperation, broadening the ground for a deeper examination of incentives to deviate from the former practices through a 'backstabbing-like' defection.

2.1.4 Benefits of faithfulness – a liberal and legalist view of alliances

Much has so far been said about systemic incentives and obstacles states encounter entering and maintaining alliances, but mostly from an epistemological standpoint of negative interests and survival – of both states and their politico-security coalitions. However, another point of view is needed to show why states pursue alliances under and despite the well-known conditions of international anarchy and all the uncertainty it brings along. Do states ever stop pursuing their relative gains for the sake of collective well-being? What is the central process of international alliances' life cycle and how does it relate to a particular state's security? What is that states do *jointly* to promote their security? Answering these questions will, as outlined in the definition of our approach, comprise a conceptual startup, which is consequently developed into a

[32] Brought into the model in the previous section

broader explanation. Simultaneous identification of fundamental explanatory notions will, as previously established, be followed.

Searching for an answer of the formerly posed questions demands a comprehensive overview of IR theories. Nonetheless, we will condense this abundant work by simply starting from the most general concept that broadly embraces all the diverse structures, practices and processes that constitute the substance of international alliances – the concept of *cooperation*[33]. Apart from its generality and inclusiveness (that provide us with a possibility of seeking specific features of intra-alliance dynamics in an extended and better grounded way), this concept will also help us shed some light on somewhat more optimistic and collaboration-marked aspects of international alliances. In order to deploy it in our study, we will:

1. *Define* what *cooperation* is and what it is not in terms of intra-alliance relations.
2. Discuss what kind of organizing principle cooperation entails in international relations, i.e. we will *relate this process with its structure*, *international institutions*;
3. Explain the *benefits of international institutions*, as it will show why states, in spite of their concerns for relative gains of their potential allies, do decide to engage in international politico-security alliances;
4. *Critically apply* this general framework to the case of intra-alliance relations in order to incorporate it into the meta-model. This input will be framed as five general points drawn from the entire discussion that follows.

The course of analysis will inevitably touch upon 'the other' confronted 'neo'-camp of IR – Neoliberal institutionalism, as most of its argumentation offers insights into the matter of the questions we posed.

The first logical step at this point is to ask *what cooperation in terms of intra-alliance relations actually means*. Keeping the realist(ic) course of analysis, we will purposefully target Robert Keohane's definition of cooperation, as it will smoothly open our attempt to bridge the gap between the realist and liberal assumptions on international cooperation. This is because Keohane straightforwardly defines interstate (or intergovernmental) cooperation as a situation that " […] takes place when the policies

[33] On specific strategies and courses of cooperation, see chapter no. 3

actually followed by one government are regarded by its partners as facilitating realization of their own objectives, as the result of process of policy coordination" (Keohane, 1984, p. 51). Analytically construing the matter of this definition, we identify its composing elements as:

- Cooperation is a process whose subjects (herein "governments") are parties consciously involved with it;
- Objects of cooperation are those parties' respective policies;
- Each party's objectives are cooperation inputs, i.e. they are *a priori* given to cooperation;
- Cooperation is conditioned by a mutual recognition of benefits seen as facilitated accomplishment of individual objectives of parties involved;
- Benefits are gained through a process of policy coordination, which means that cooperation entails policy adjustments and, conversely, objectives accommodation[34].

A conclusion that can be aptly drawn is that all major elements of this definition congruently match the relevant neorealist views, except for one: cooperation produces authentic, distinguishable and independent effects on international relations. It can actually transform the interests of the parties involved, which renders valid qualifying it as a separate factor of impact in international relations. The fact that states are engaged in a common process influences their perspectives and perceptions[35], therefore the process of foreign policies creation as well. Understood in this way, formations hosting cooperation fit even Waltz's definition of structure (Waltz, 1979, p. 73). It would be, still, too daring to claim that Keohane managed to actually bridge the gap between the camps[36] by framing such a broad definition of cooperation. What is more justifiable to claim is that this definition provides a possibility of a reasonable inclusion of neorealist arguments into the meta-model, making it theoretically 'bipartisan' as approximate to neutral, and coherent at the same time.

[34] Here, however, the notion of equifinality has to be taken into account as well (Croft, 1996)
[35] Given that states "regard" policies pursued by another government as facilitating accomplishment of their own goals.
[36] Although the gap itself is far narrower than usually seen, as admitted by Keohane himself (1998, p, 86)

Another step towards a comprehensive theory of international cooperation is establishing the concept's *differentia specifica* and distinguishing it from its pre-World War II idealist roots, i.e. from the notion of harmony. Having defined harmony as a "situation in which actors' policies (pursued in their own self-interest without regard for others) *automatically*[37] facilitate the attainment of others' goals" (Keohane, 1984, p. 49), Keohane acknowledges the initial state of objectives or interests divergence that is to be decreased only through a process of cooperation (Ibid.), meant to foster negotiations and policy coordination. Another strand of algorithm of initial negotiations as a prelude to cooperation is the point of discord, i.e. a situation in which actors' policies did not become significantly or critically compatible with one another. This is an additional theoretical *rapprochement* with realism, as Keohane acknowledges that disabling disagreements can be an outcome of intended cooperation. Moreover, discord is not value-free by any means, it represents a negative tendency in international relations with an inherent disagreement potential in the future, being "a situation in which governments regard each others' policies as hindering the attainment of their goals, and hold each other responsible for these constraints" (p. 52). Acknowledging a potential for conflicts preserved in the practice of cooperation renders this theory much closer to everyday reality of international institutions abounding with disputes and disagreements. On the other hand, Keohane further supports the construction of cooperation-conflict relation bluntly putting that cooperation "should not be viewed as the absence of conflict, but rather as a reaction to conflict or potential conflict" (p. 53).

Defining the substance of the term furthermore, Keohane emphasizes two crucial aspects of cooperation:

- It is a *pattern of behavior*;
- The quality of behavior that constitutes the process of cooperation is *altered* over time (1982, p. 52).

This particularly means that author assigns the attributes of structure-building potential to the process of cooperation[38], and, accordingly, a capacity of that process to perform

[37] Emphasis K.R.
[38] Think of the example of *Conference on* Security and Cooperation in Europe that grew to become the *Organization for* Security and Cooperation in Europe.

genuine and independent influence on actors involved, which is so vehemently disputed by neorealists, as previously shown.

Building on top of the former point, Keohane defines the *structure* that induces and is induced by the *process* of cooperation: international institutions, broadly seen as international regimes, i.e. concept far wider than the reality of international intergovernmental organizations. First, he overtakes the argumentation of John Ruggie, who defines regimes as "a set of mutual expectations, rules and regulations, plans, organizational energies and financial commitments, which have been accepted by a group of states" (1975, p. 570), and the definition proposed by Stephen D. Krasner, particularly stating international regimes are

> sets of implicit or explicit principles, norms, rules and decision-making procedures around which actors' expectations converge in a given area of international relations. Principles are beliefs of fact, causation, and rectitude. Norms are standards of behavior defined in terms of rights and obligations. Rules are specific prescriptions or proscriptions for action. Decision-making procedures are prevailing practices for making and implementing collective choice. (Krasner, 1983, p. 2)

What Keohane keeps in his definition is a view on international regimes as social institutions. Although a significant part of his argumentation is disambiguation of norms, rules and principles, we will not go further into it, as it has no practical implications for our meta-model. What is relevant for our study is to emphasize that norms, rules and principles of international regimes are interrelated through a common reference to sanctions about behavior, as they define certain actions as prescribable (allowed or explicitly ordered and expected) or proscribable (liable to banning), thus implying responsibilities and obligations, even when they are "not enforceable through a hierarchical legal system" (Keohane, 1984, p. 55), as it is often a case with international alliances. What these injunctions provide is a framework that is meant to foster and facilitate the spread of cooperative agreements among states. Still, Keohane openly recognizes that the efficiency of those structures largely depends on nation states' will, sublimed in the idea and practice of sovereignty, which, consequently, means that international regimes should not be regarded as components of some world order beyond the nation-state. On the contrary, the only realistic view is the one that recognizes that regimes are "arrangements motivated by self-interests" (p. 59). However, those

arrangements contain mutual commitments of the states involved, thus developing the according mutual expectations. These latter are in turn boosted by states practicing rules, norms, principles and decision-making procedures outlined in the core of the regime. Pursuing this track finally brings about authentic influence of the process of cooperation on its participants, discernible from a mere aggregation of individual national interests.

To fully prepare the ground for a specific examination of international politico-security alliances as international institutions (regimes), we should also pose the question of *benefits* that *international regimes* bring about. How do they advance individual and collective interests worldwide, if so? What are the benefits of believing in the grand promise of international institutions, albeit false? What is that promise actually all about?

As widely argued, international institutions increase, but are also a manifestation of a growing international interdependence. It means they offer a context for interactions of participating states and sub-state or trans-state subjects. However, in terms of international relations, this interacting context possesses formal and informal channels of diversified and multi-track communication, most broadly defined as an exchange of information, whatever their ontology and nature might be. Bearing this in mind, Keohane generalized benefits that arise from this communication as:

- *Reduction* of transaction costs (costs of making and enforcing agreements) by using regimes' multilateral structure and standard communication channels;
- Providing states involved with *reciprocity* (which entails an option of institutionalized retaliation and retribution of costs) i.e. incentives for governments to cooperate and keep their own commitments so they can expect others to do so as well;
- Promoting *transparency* supported by procedures (dealing with a class of problems in a standard and expected way, systematically monitoring states' compliance with the obligations and commitments they made) and encouraging international negotiations and agreements, thus decreasing the lack of interstate trust and uncertainty;
- Creating a sense of *continuity* and *certainty* in international relations;
- Producing a *synergetic, net effect* from multilaterally joint capacities, which acting together overweighs their individual effects aggregated.

- Once engaged in multilateral structures, it is *difficult for a particular state to calculate its relative gains*, and possibly withdraw from cooperation, as the channels of benefits get multiplied with the channels, quality and amount of information state receives (1998, pp. 85-87).

This argumentation on benefits relies on a substantially simple mechanism: the more we know about each other, the safer we feel. The safer we feel, the less are incentives we have for defection and cheating. Compatible, although more specific approach can also be tracked in the work of Bearce, Flanagan and Floros on the inverse proportionality of the intensity of information flow within alliances and propensity and frequency of military conflicts among member states. They use the results of bargaining models of war to show that an insufficiency of information about other states' relative military capabilities functions as a salient cause of war. Taking a step further, they argue that such international institutions as alliances do provide its members with that sort of knowledge, thus decreasing both the sense of uncertainty and problems that arise due to a reoccurring lack of information. Their extended argumentation supported by vast statistical data and analysis leads to a conclusion that capability distribution has no effect on dyadic military conflict when a pair of states participates in an institution of close cooperation. Based on this conclusion, they argue to have proved that military alliances do produce independent effects on interstate relations, due to a genuine process of information exchange that they embrace (Bearce, Floros, Flanagan, 2006). Given that we have already abstracted from defensive Neorealism's myopic argumentation on the purely derived nature of international institutions, remarks made by Bearce, Flanagan and Floros can be straightforwardly incorporated into our meta-model, with a main point on the significance of *intra-alliance information exchange*.

Additionally, remarks far more optimistic than the realist ones have been made on the notion of trust, i.e. on the standpoint it should be regarded from. As Robert Axelrod confidently argued, the essence of cooperation is a widely flawed conceptual point, as its core is not really trust, but a specific relation's history, or as he put it "the durability of the relationship". This moderately optimistic view advocates that actors can, under "right conditions", come to cooperate with each other on a basis of trial-and-error learning about the options of reciprocal rewarding. The process would prospectively lead to an

established knowledge of how the cooperation can be improved and what sort of behavior is to be avoided for it to succeed (Axelrod, 1984; in Williams, Goldstein, Shafritz (Eds.), 2006, p. 334). Policy recommendation derived from this argument is basically never to defect first and to seriously value a long-term continuation of cooperation even if the other side has defected once, which still does not exclude an option of short-term retaliations[39].

However, Keohane himself does not reject the notion that states cheat not only because they act preventively or retributively, but also because they simply see their short-term benefits largely multiplying by breaking the chain of structural trust and commitments, i.e. by selfishly misusing the total amount of intra-institutional information flow. This kind of intra-institutional behavior obviously rises from a pronounced valuation of relative gains, especially if the global constellation of power follows the bipolar lines of delineation (Keohane, 1998, p. 83). For these reasons, an unconditional cooperation, even with the 'reassuring provisions' of regimes, remains a risky endeavor in the international arena.

Instead of discussing the validity of Keohane's arguments on benefits of international regimes, we will first *apply this general model* to the case of this study's concern: international politico-security alliances. The assessment of this argumentation will be thoroughly provided once the game theory input is framed, as its main purpose will be to depict incentives for actors' particular modes of behavior in more details.

Firstly, concerning Keohane's definition of cooperation, it can be confidently concluded that the conditions of its applicability to the case of alliances are fulfilled. As the author himself put it quite clearly, "alliances are *institutions*[40]" (Keohane, 1989, p. 5) As for the approach of this paper, a deducible conclusion is that the elements of the definition applied widely encompass the phenomenon of cooperative processes within international alliances, in terms of their *actors-participants* (allying states), *objects of cooperation* (individual states' policies), *cooperation inputs* (individual allies' given objectives), *conditions for cooperation* (allies' mutual recognition of shared benefits) and fundamental *functioning mechanism* (process of policy coordination). Still, this study

[39] See chapter no. 3 for the point on non-myopic equilibrium
[40] Emphasis in the original

adopts the notion that these points define an ideal-type international politico-security alliance. Practical cases often show functioning deficiencies[41] and necessities for upgrades[42] with various regards of this definition.

Secondly, international politico-security alliances do abound with allies' differences in opinions on specific issues, even conflicts of various scales, as our case-study examination will show. Thus, they contain both interests converging and diverging capacities.

Thirdly, though a process of intra-alliance cooperation, coordination and negotiations, patterns of behavior are formed and maintained, entailing a reproduction of the aforementioned capacities for interests' convergence/divergence. Allies may solve or not specific issues among themselves, but modes of negotiations that maintain an alliance are developed into a pattern over time. This pattern, this accumulated coordination, bargaining and negotiations (or a lack of them) is what forms an institutional history of alliances, i.e. a *de facto* institution in a broader sense[43]. With respect to international alliances, examining patterns of intra-alliance politico-security support (i.e. its frequency and distribution) can be a particularly insightful project, prospectively operationalizing and deepening the applicability of the meta-model this study is aiming to construct.

Fourthly, alliances do encompass a vivid field of principles, rules, norms and decision-making procedures. However, depending on a nature of intra-alliance relations, their intensity and extensity, as well as the strength of interests that underlies them, alliances rest upon varying grounds of 'connective regulatory tissue'. Some of them, based on overarching values just as interests, place their focus on principles and norms (i.e. the evaluative dimension of alliances), whereas the ones strongly focusing on particularly bonding interests develop more detailed decision-making procedures, as the

[41] Cooperation inputs can be undemocratically or non-procedurally determined by particular actors, thus complicating functioning of specific alliance (how stable ally of the US is the Pakistani president Pervez Musharraf, or can this definition legitimately subsume modern Afghanistan?)

[42] Some allies, like the EU, are placed above the national, i.e. at the supranational level, which affects the functioning mechanisms and nature of its alliances (although NATO members are EU's Member States, not the EU as such, *de facto* transformation of NATO's core into an EU-US alliance is induced by the European CFSP). Also, the history of International Relations keeps record of politico-security alliances of a rather mixed type: among state and non-state actors (US-Taliban alliance in Afghanistan during the 1980s, NATO-UÇK alliance in Kosovo in the 1998-1999 period)

[43] Example of the UK-USSR permanent tensions during the World War II is a striking one, but examples of various regular intra-alliance consultation practices prove the case as well (e.g. EU, African Union, Franco-British, Franco-German summits etc.), although in a more general way.

values that could connect them with more certainty are remotely expressed and weaker, if any (take the example of the Little Entente formed by Yugoslavia, Czechoslovakia and Romania in 1920-192, whose main purpose was to contain the revisionism of Hungary in the aftermath of the World War I)[44]. Whatever the prevailing nature of these commitments and their subsequent expressions, it is crucial to emphasize at this point that alliances do need an underlying *accord*, a fundament of their existence, be it formal or informal, written or unwritten, *expressis verbis* or tacit. This accord, whatever its nature, contains this core exchange of commitments and mutual expectations, rights and obligations that jointly frame a cooperative nucleus of the alliance in question, being an immediate[45], or perhaps an indirect[46] emanation of states' individual interests. As the most general level of understanding among alliance members, this essential and primary accord conditions the state of subsequent specific rules and procedures. The accord can take a form of an official international legal document, i.e. of a written and subsequently ratified agreement, as in the case of NATO's Washington Treaty (1949), or can remain at the level of a mutual tacit or explicit understanding (as in the case of the US-Iranian alliance before the Islamic Revolution of 1978), without ever acquiring such a degree of codification[47]. The reason for insisting on the essential exchange of reciprocal rights and obligations, and not on the form of such committing is that both preserve the same legal value. As this idea of accord is what makes an alliance come into practice and fulfill its purposes, this study adopts this broader view of its legal framework, concentrating rather on the content of commitments, mutual expectations and their implications. However, as the form is often the essence at the same time (especially in diplomacy), the way of defining and determining these rights and obligations, and possible underlying values is

[44] The fact that a formal agreement legally establishing this alliance was signed only in 1933, straightforwardly confirms this notion, as at this point the informal and genuine commitments started loosening as the pressure of revisionism started increasing. Therefore, the commitments (in order to remain credible) needed a legal support. On the nature of Little Entente's legal grounding, see Galitzi, C. (1933). The Balkan Federation. *The Annals of the American Academy of Political and Social Science,* Vol. 168, No. 1, 178-182; and Leeds, B. A. (2003). Alliance Reliability in Times of War. *International Organization*, 57, 801-827

[45] If declared interests and goals match the actual ones

[46] If the actual interests and goals are of a rather clandestine nature

[47] Although the practice is to label these *informal alliances* as '*alignments*', we will for the sake of simplicity use only the former term, given that our meta-model aims at logically encompassing them both. Moreover, this differentiation has no implications on those international coalitions' legal validity (but does have ones on their legal certainty), which is an additional argument for this study to use this single term.

highly relevant, especially when differences in interpretation of those arise. Closing this remark, we will conclude that the importance of this founding accord lies in its content and the consequent expectations arising from it.

Fifthly, whether international politico-security alliances actually deliver expected benefits to participating actors is a question of specific case studies. What is argued by Keohane and other relevant authors can hardly be rebutted, but whether the cooperative capacities and incentives or the defecting, or even conflictual ones prevail within alliances is what our meta-model focuses on by building up a general examination framework. Edifying a structure that can constructively test practical developments of these potentials remains this paper's ultimate goal.

2.1.5 Further conditions facilitating the functioning of alliances

A milestone question of any general theory of international alliances is whether entities other than compatible interests and goals (whatever their nature is, thus including both 'neo' camps and their intermediates) condition the functioning of alliances and, more specifically facilitate their formation and maintenance.

By far the most examined factor is the regime type of alliance members, and its relation with the logic of alliance behavior. Sources of these discussions can be traced back to a vast field of theories of IR, particularly the theories of constructivism, economic interdependence, mutual commitments and democratic peace theory. However, the seminal work of Karl Deutsch on security communities has for a long time been an umbrella theory for a number of concrete inquiries. Herein, we will shortly exhibit his definition of security community as a group of people (and consequently induced group of states) that managed to become "integrated", which specifically means attaining a sense of security, institutions and practices widespread sufficiently to assure those actors' mutually dependable expectations about their pacific relations. Herein, community means *belief* that common problems are resolved through "peaceful change" (Deutsch, 1957). Developing the case furthermore, we conclude that within a security community war and large-scale physical force are conceptually excluded as a means of dispute settlement, and that the community in question possesses institutional arrangements to resolve

differences in a pacific way. What delivers this sort of community is an integration process, understood as a dialectical flow, within which the experience of security entails the creation of institution, which, in turn, fosters security. This dialectical process is furthermore rendered possible due to shared values, norms and beliefs among the community members. Two types of security communities have been historically recorded: amalgamated and non-amalgamated or *pluralistic*. Looking at their specific features, as defined by Deutsch[48], we see that the case of international alliances could rather be related to the concept of pluralistic communities, as the amalgamated ones are generally actualized in unitary societies and modern states. On the other hand, pluralistic communities have historically been easier to establish and later on maintain. Explicit conditions for their formation and further development are by far less demanding, including the compatibility of crucial political values, responsiveness to each other's needs and messages, and a partial mutual predictability presupposing a stock of knowledge that actors potentially have about each other (Ibid.).

Question before us is whether this deep interconnectedness and interstate altruism stimulate a creation of endurable and stable alliances. Although bringing up specific examples, Deutsch cannot really help us answer this question, due to a largely abstract level of his argumentation if compared with rather specific needs of our study. Nonetheless, specific investigations drawing their inspiration from this and other grand theories can try to give us an answer.

Therefore, we will introduce a more focused study equipped with quantitative methodological instruments to verify a general hypothesis previously implied: Are states of similar regime types[49] more likely to ally with each other[50]? An inquiry conducted by

[48] Conditions for amalgamated communities include:
1. the mutual compatibility of crucial values;
2. a distinctive way of life;
3. capacities and processes of cross-cutting communication;
4. high geographic and social mobility;
5. multiplicity and balance of transactions;
6. significant frequency of some interchange in group roles;
7. widening of the political elite;
8. and high political and administrative capabilities
9. willingness and ability of majority of the politically relevant strata to constructively manage common institutions by expanding and fostering them (Deutsch, 1957)

[49] Regime type in relation to a state's ideological orientation is taken as an operationalization of "compatibility of crucial political values" (Deutsch, 1957)

Lai and Reiter (2000) generally confirms this hypothesis, based on a fruitful examination of all pairs of states between 1816 and1992. However, their conclusion is related only to the post-World War II period, i.e. the age of Cold War, when the ideological polarization was vehement and forceful. Statistical population taken into account within the temporal framework concluding at the point of World War II cannot confirm the hypothesis. Another dilemma before us is whether the systemic conditions since 1992 have changed (and they certainly have) to the extent that would render these conclusions irrelevant for the nowadays alliances. We should not forget the widespread doubts about the survival of such an alliance as NATO in the aftermath of the Cold War[51]. As we see, it did manage to survive and overcome the obstacle of uncertainty about the future scope and purposes, but if that happened due to essentially unchanged systemic conditions or due to a reformed and adjusted political project is an imposing question. Even if the conditions are essentially changed, did ideology and type of regime stop playing a role along the way? For this to answer we will need a new set of extensive studies, at the moment inaccessible, due to a fact that the current state of international affairs poses an immense variety and intensity of systemic changes and transformations. However, what we can say for sure is that the major ideologically driven security alliance of the modern world, NATO, still exists and its normative nature of existence is repeatedly stated. The same is true for the EU, whose normative foundations are even more pronounced, and we do have to recognize a security dimension of its structure. All the aforesaid allows us to say that ideological closeness and similarity (along with the regime type[52]), as an expression of shared values, norms and beliefs still plays a considerable role, although a firm causation cannot be established. This means we will take this factor into consideration while constructing the meta-model, though with a due scientific caution and critical stance.

[50] At this point, a reference to Kant's Perpetual Peace (1795) serves us to show original springs of this strand of researches, but including it into our discussion is impractical, due to only circumstantial, rather philosophical and quite broad links with our study-object.
[51] Kenneth Waltz (1993, p. 76) predicted a soon demise of NATO, saying "NATO's days are not numbered, but its years are".
[52] Concepts are interrelated, but are not matching completely

2.2 Completing the meta-modeling input of theories of alliances

Based upon the previously exhibited argumentation and vast theoretical material, we can, closing this chapter of our paper, finally frame the first portion of inputs that will prospectively lead to a constitution of a general model meant to examine whether a constructive and convincing analysis of intra-alliance relations can be conducted through a congruent combination of two courses of theories: theories of alliances and game theory. As deduced, this input includes all the elements abstracted from the theories and models decomposed and reconstructed to form a coherent meta-theoretical body. It finally provides us with a possibility to make a structure (later to be combined with properties of game theory) of *elements (components)* and *determinants* of international alliances. However, to keep the self-critical and scientifically constructive orientation of this study, we will present some quandaries and methodological dilemmas related to the abstractions we made, suggesting their possible examination, testing and correction in future, possibly more operationalizing studies.

We first enlist the abstracted components of international alliances derived from the earlier discussion. Those are as follows:

1. Alliance members – systemic actors;
2. Fundamental and underlying alliance accord (agreement, contract) with (possibly) issues, common goals, principles, alliance nature, temporal and spatial frame of alliance defined;
3. Power flow and capabilities distribution within alliance;
4. Allies' common and compatible interests of a diverse nature[53] (individual, collective, security or autonomy-driven);
5. Intra-alliance cooperation as coordination (information exchange etc.)

Methodological and theoretical quandaries related to these components are:

- How to ontologically categorize the constituting components of alliances?
- How to systematically follow their life-cycle?
- Should criteria for a construction of 'solid' and 'survivable' components be defined, and if yes, where to draw normative benchmarks from?

[53] also appear as a determinant

- How to assess each component's influence on the state and development of alliances?
- Should a normative stand be taken while performing an act of assessment?

Determinants of international politico-security alliances as abstracted from the preceding discussion on a set of theories and models are:

1. *Intra-systemic stimuli*
    - Allies' varying interests and motivation to ally with other actors;
    - Allies' particular definitions of national interest;
    - Individual member's internal politics, socio-economic, demographic and military conditions, political and regime similarity of allies;
    - Relative significance of issues with converging/diverging interests among allies, resulting parallelogram of allies' diverse converging/diverging interests;
    - Allies' relative capabilities;
    - Allies' strategies of intra-alliance behavior;
    - Intensity, extensity and density of information flow among allies,
    - Allies' individual and common historical and security legacy.
2. *Extra-systemic (environment's) stimuli*:
    - Number of extra-alliance actors;
    - Fundamental and situational set of military, political and economic challenges, risks and threats – all regarded as stimuli of an intrinsically security nature (whether hard or soft), in accordance with essentially securitized epistemology of interstate alliances.

Theoretical and methodological quandaries springing out from this structure are identifiable as well:

- How to discover what, who, to what extent and how determines the life of interstate alliances in their both static and dynamic dimension?
- How to measure the intensity (individual and combined) and direction of the determinants' influence, i.e. how to vector it adequately?
- How to discern the impact of fundamental and situational determinants?

- How to calculate and predict the effects of the determinants in action?
- Is it possible to manipulate the determinants?

The following section will deal with an examination, deconstruction and re-composition of game theory (or theories) in order to make it coherent with the current structure of the meta-model.

3. ADAPTING GAME THEORY: FRAMING INPUT2

This chapter's task will be to critically examine the logical structure of game theory in the light of its compatibility with theories of alliances so far reconstructed into a joint model. Requirements of mutual logical and substantial coherence will remain two crucial criteria when joining these two major theoretical strands placed in the focus of this study. An overall complementarity of these two theoretical clusters is provided through a general applicability of game theory as an analytical instrument and tool of a rather formal epistemological role, whereas the previously discussed theories of alliances present the very matter of this study, i.e. its ontological essence.

3.1 Postulating game theoretical methodological apparatus in terms of intra-alliance relations

This chapter is opened by an attempt to initially integrate the fundamental structure of game theory with the input1 formerly derived from theories of alliances. However, this integration will comprise double-track adjustments of both theories and their tools, although game theory is here defined as a formal structural component. This approach will provide us with more knowledge on how far we can go with a study of intra-alliance relations using quantitative methods. The incompatible aspects of both theories' inputs will be used to draw consequential conclusions on the limitations of the approach[54].

Conventionally described, game theory features several definition elements:

1. *Players* (herein, alliance members);
2. Set of *possible actions* at alliance members' disposal;
3. *Information available* to alliance members when choosing their actions;
4. *Payoff consequences* for each alliance member and for every possible combination of actions chosen by all alliance members participating in the game;
5. A description of *preferences over payoffs* of all alliance members (Guillermo, 2001).

[54] Refers to analyzing intra-alliance relations with game theory apparatus

Each of these elements will be briefly discussed in the light of demands posed by internal structure, contextual positioning and nature of alliances and intra-alliance relations.

3.1.1 Players

Although this paper deals with intra-alliance relations and takes international politico-security alliances as its targeted environment with a set of specific and general features, we can hardly deny that an alliance is only a part of a broader context of international politics, that its members do not have isolated foreign policies dedicated to concrete alliances solely, and that various extra-alliance stimuli affect its development and internal relations, which is explicitly recognized in this study once we included "extra-systemic" stimuli as a group of determinants. International politics are dynamic and with widely interdependent actors and events. The fact that intra-alliance relations occur within the structure and environment we are primarily focusing on does not mean excluding a broader context of analysis, whatever the stance is accepted on alliances' capacity to produce independent effects on actors, their relations and international system – Neorealist or Liberal Institutionalists, as previously discussed. Alliances are formed to deal with *others* outside of them, at least declaratively.

In terms of game theory and its application, it means that a game, although focusing on alliance members, can include extra-alliance actors as well, be it just at the level of a sequence of game (if the game is sequential) or the whole of it (sequential, simultaneous or iterated). These actors will therefore not be regarded as intra-alliance ones, but their inclusion will mean they substantially influence the state of relations among allies[55]. Thus, the criterion for inclusion of this class of actors at any point of the game is their considerable influence over other players' strategic choices and over the outcome of the game.

[55] As in the Gambian-Senegalese alliance that reached its summit in 1982, when the Treaty of Confederation was signed. In the case of this two-member alliance, the UK and France, former colonial rulers, traditionally had a major say.

3.1.2 Set of possible actions: strategies and tactics

Given their highest level of generality and long-term orientation within a game, strategies[56] represent an umbrella element of games, bringing together their static and dynamic perspectives. Some of specific strategies at alliance members' disposal have already been discussed in the previous section. However, as defined at that point, those strategies are rendered highly specific and contextual once fit in the game theory framework. Game theory brings a fruitful generality of strategies designation easily applicable to a variety of cases with only a slight adjustment needed. Specific modes of games (prisoner's dilemma, chicken game, coordination and assurance game, trust game etc.) have naturally concrete set of actions and the according usual labels (such as 'to cooperate' and 'to defect' in prisoner's dilemma), but their fruitfulness is precisely in their wide usability. Moreover, all these strategies are rather ideal types than clear-cut empirical cases, which means their applicability and adjustment potentials are indeed wide[57].

Two notions need to be discussed here particularly. Firstly, to examine strategies in game theory means to place our focus on the process of their interactive determination by players. Applied to our case, this process is manifested through dynamic practices of decision-making, where the choice of a concrete strategy by an alliance member is tightly

[56] More precisely, possible actions become strategies only when rationally, intentionally and purposefully considered by a player, i.e. a strategy presupposes an active attitude of a player towards actions it can take

[57] Consider an example of alliance A, with a {a, b} dyad, where allies are playing a coordination game. They are conducting double-track bilateral negotiations within an international organization over an issue. The allies are supposed to divide the tasks, and each of them is to lobby one out of two states which they believe need convincing in order to accept their motion. If they insufficiently coordinate their actions and both choose the same state to lobby, they fail to accomplish their goal. Standard strategies are as sketched in the matrix (Figure 2) with two Nash equilibria. However, the allies might have compatible preferences over who to lobby and they might both choose the same Nash equilibrium (Figure 3), which is known as 'pure coordination game'. This differentiation in strategies comes precisely from an array of practical cases.

/	1	2
2	**10, 10**	0, 0
1	0, 0	**10, 10**

Figure 2

/	1	2
2	***10, 10***	0,0
1	0, 0	**5, 5**

Figure 3

co-determined by strategies other members have at their disposal. A probability estimation of whether certain strategy will be played by another ally has a salient role. In this sense, strategies in game theory are regularly dichotomized as the *pure* and *mixed* ones. Pure strategy concept refers to a situation where a player who is given a set of available strategies chooses to take an action with probability 1 (100%). Think, for instance, of the example where allies engaged in a military conflict against a third party have to perform attacks on two crucial enemy fortifications. The coordination problem posed is as following: if both allies choose the same fortification, they fail to accomplish the goal, and *vice versa*. If ally A is playing strategy S_1 (which designates attacking the fortification P) with probability 1, then it is said to be playing a pure strategy. However, if ally B's strategy is conditioned by another event[58], i.e. the outcome of that event effects the strategy choice, then the probability of ally B choosing one of the two strategies (attacking either fortification P or Q) is not 1, and it is said to be playing mixed strategy. For instance, if in an alliance of unequal two states often choose to free ride when it comes to investing in joint military capabilities, but one of them does it all the time and the other one does it only if its national economy is declining, the former is said to be playing a pure strategy and the latter a mixed one, as the probability of actually deploying free-riding is not 1. Game theory simplifies the choice of a specific mixed strategy to a random pick[59]; however, the above depicted cross-conditioning can always broaden and complement such a simplistic view. Subsuming specific intra-alliance strategies[60] as either pure or mixed[61] is crucial, as it implies the possibility of identifying pure or mixed strategy Nash equilibrium. This in turn affects the strategic game planning, as predicting the best response in the course of a game remains the ultimate goal[62].

Secondly, we should tentatively differentiate between:

[58] Might be coin flipping, waiting for the outcome of a nearby battle etc.
[59] As with using the coin flipping
[60] E.g. free-riding, buck-passing, chain-ganging or bandwagoning
[61] Consider an example of intra-alliance relations within NATO right before the war in Iraq in 2003. Therein, the UK decided to almost unconditionally support the US in its decision to intervene, so it could be said the UK was playing a pure strategy, whereas the decision of Belgium regarding the same issue largely depended on whether the EU could come up with a unanimous position; therefore, Belgium was playing a mixed strategy.
[62] Ally who is capable of predicting a Nash equilibrium has, for example, a considerable negotiation and bargaining advantage, as it can take a lead by suggesting credible and plausible solutions with a rational grounding.

1. Strategies allies play in order to maximize their respective gains in intra-alliance games[63];
2. Strategies played in games that involve extra-alliance players, such as third states[64].

Still, we should warn that throughout the game a substantially connected sequence of strategies can belong to different contexts, for instance: a strategy played in relation to an ally can entail further extra-alliance strategies etc. Take, for instance, the example of Tito's Yugoslavia's 'divorce' from the Soviet Union in 1948, which entailed a change of the US politico-security strategy in Europe, opening the opportunity for a rapprochement with communist Yugoslavia and increased influence behind the 'iron curtain' (Young & Kent, 2004, pp. 82-83). This observation stands strongly on our conclusions from chapter 2, where we explicitly outlined extra-alliance stimuli to be one of the determinants of intra-alliance relations.

Let us now consider a short term dimension and focus of games by touching upon the role of tactics. A way of analyzing games that involve allies at least ostensibly[65] engaged in the same project is focusing on specific moves they intend to play and see how they transform their strategic planning into actually played strategies and, more specifically, tactics[66]. However, as the purpose of this meta-model is not to strictly encompass a detailed game theoretical framework, but to set a methodologically plausible research outline, we will only generally suggest how some methods can be used in our analysis. Apart from that, our meta-model has a severe requirement of staying flexible

[63] Herein, certain alliance appears as an environment, a context within which a game is played, i.e. as a *structure*

[64] In this case, alliance will be taken as an actor in the game, given that decisions are commonly taken by all allies and reached through a process of collective, intra-alliance goals and actions accommodating, i.e. it will be treated as an *agent*. However, it is possible to have a game in which extra-alliance players are involved, but the alliance itself does not play a suggested role of an actor, as decisions are not being taken at the collective level, or are not being taken at all (e.g. the issue of Iraq, an active player before the invasion in 2003, divided the Transatlantic alliance, thus impeding NATO-level decisions; this, nevertheless, did not make a discussion on Transatlantic alliance obsolete – quite contrary).

[65] That is a minimal condition to discuss relations among allies and not just relations among any given competing states or state-like entities.

[66] Drawing a distinction between strategies and tactics always represents a challenging defining effort. Herein, we will define tactics as a set of means allies deploy to achieve particular and short-term ends as a part of rather mid and long term prospects of actions defined as strategies (Bacharach & Lawler, 1981, p. 234).

towards a broader apparatus of game theory, which means including its features in specific, rather operationalizing and more empirical researches must be allowed.

Focusing on move-to-move[67] planning that allies do can potentially lead to an infinite and futile examination of a series of best response and alternative strategies that might be developed in the course of a game. However, setting a general framework is a strict scientific imperative. Therefore, we suggest a general application of two methods that allow for further operationalizing: *searching technique* and *evaluation function*. Facilitating a focus on a move-to-move plane of the game, these methods offer insights into allies' tactical, not only strategic dimension of planning, which is why they are jointly known as parts of the *tactical analysis engine* (Winston, 1992, p. 95).

The main purpose of the search technique is to determine which class of moves is to be considered as a response or maybe a first-shot tactics, and returns the move it considers best at the end of the search. Following this technique, a tree of moves and counter moves is created, and the 'leaves' of the tree are assessed according to an evaluation function. The evaluation function, in turn, reports how good a position looks to the player whose move it is. By far the most used combination of the two methods is so called "mini-max" technique (Katz & Butler, 1994), which assigns to prior strategies a range of minimum/maximum value of their posteriors, and, by using inverse rationality[68] it maximizes the root of the created tree, i.e. the original strategy.

3.1.3 Information available to alliance members

The previous chapter has in dept shown why a regular and controllable information flow among allies is crucial for a survival and well-functioning of an alliance. At the same time, it has clarified how alliances precisely through an information share and exchange among the members involved reduce the possibility of wars[69]. The point of disagreement was whether expectations on regularity of such a practice should evolve, where Liberal Institutionalists answer this question affirmatively, and Classical Realists and Neorealists, on the other hand, quite promptly negatively. Building trust and

[67] Or shot-to-shot
[68] i.e. backward induction
[69] See Chapter 2, section 2.1.4

reducing the possibility of cheating[70] is, no matter what, a question of adequate arrangement, and we have clearly, by joining two opposite logical operations of generalization and abstraction, reached a model of alliance institutionally equipped with provisions on information exchange, precisely due to its importance for intra-alliance relations.

In terms of game theory, information stock basically refers to knowledge that allies posses about each other's possible strategies, payoffs and strategies actually played. Information distribution can be structured along the following lines:

- Information on each other's strategies and payoffs (*complete information*);
    - Information on each other's strategies actually played (*perfect information*, possessed only in sequential games);
    - Information on all the possible outcomes/payoffs, but not on the strategies actually played by other players (*imperfect information* in simultaneous or *de facto* simultaneous games)
- At certain stage of the game (i.e. its specific node), a player whose turn it is to 'shoot' knows less then a player who has already made its move (*incomplete information*).

Still, allies' struggle to know as much as possible about each other goes far beyond this and includes an immense field of desired information, including their foreign policy declared (manifested) and actual (latent) interests and goals, state of their national economies, structure and intensity of their international trade (especially their main trade partners), domestic politics, military capabilities, defense budgets and planned investments, public and clandestine international accords, all the way to cultural issues and information about allies' leaders' individual characteristics and private life. This is obviously a range of information impossible to include all at once in game theory-based researches and subsequent calculations. However, all these information make a unifiable stock of knowledge that is being analyzed and classified according to multiple criteria of relevance, study object, reliability, practical employability etc. After information and data have been collected, processed and classified, they can be used as manipulable entities of

[70] The two processes ideally work simultaneously, they are interdependent, but are not the same: the former means 'we *do not need* to distrust each other and *we do not want* to cheat', whereas the latter means 'practically, we *cannot* cheat on each other', i.e. there is a systemic, institutional obstacle to cheating.

decomposable data sets, meaning their modes of usage can be vast, due to capacities of information recombining[71]. Limitations are posed by common deficiencies of general and specific logical operations, such as induction and deduction for instance.

Once one decides to systematically operates with data and prospectively apply game theory to an examination of international relations, the purpose of this data collection and information engineering (herein described quite basically) becomes to:

- Identify the strategies at alliance members' disposal;
- Tentatively formulate other alliance members' payoffs in the game, preferably in cardinal terms;
- Identify other members' preferences over payoffs, scale them and rank in a probable order of preference;
- Clarify the nature and sort of game that is being played;
- Predict the next move of other alliance members and infer the best response strategy in advance;
- Draw conclusion on previous, possibly unknown actions of allies using inverse rationality once the staring point, i.e. 'a safe ground' for inferring is established.

This discussion on information employability, nevertheless, is of a particular salience for our meta-modeling endeavor. It is to show how empirical data existing in the reality of everyday intra-alliance politics (not excluding extra-alliance stimuli) are transformable into a set of information relevant for a game theory consequential application. It means that an overwhelming field of information can be arranged, systematized and used to define, predict and play the game, subsuming the knowledge acquired under these three categories: information on strategies at disposal, payoffs and strategies actually chosen.

Still, a separate problem is what portion of this information set is available to players at any stage of the game? This clear-cut division into perfect and complete information insightfully identifies these insufficiencies in knowledge that players posses. Moreover, it is usually a consequence of *imperfection* and/or *incompletion of information* that players face at earlier stages of the game, or at its fundamental level. Alliance members can never know everything or possibly not even enough about their allies'

[71] Criteria of logical and substantial coherence apply here as well.

military capabilities and undeclared interests for instance. This, generally taken, lack of information, therefore remains the very source of both uncertainty and distrust, and potentials for defecting. This remark perfectly matches our observations from the previous chapter on the notion of allies' absolute and relative gains valuations as a rationalistic ground (in a reign of individual perceptions) for distrust and incentives to cheat in the context of known and 'unknown unknowns'[72], as unanticipated uncertainty and sometimes unanticipated decision variables are labeled in decision analysis (Clemen, 1996).

Generalized up to the level of game theory, this information distribution deficiency means that alliance members can hardly have perfect or even complete information that includes other allies' preferences over payoffs[73]. This claim stands for three major reasons. Firstly, allies' preference order is collective (no matter how much we analytically regard each of them as a unitary actor) and is most often drawn from a compromise of contradicting standpoints within a state, therefore, it is difficult to predict. Secondly, governments try to keep as many information as possible hidden (even in an alliance), and, thirdly, allies' preferences cannot be easily induced from their previous choices and actions.

If we try to depict this imperfection and incompletion of information schematically, and operationalize the above discussion, three major general cases can be defined. Those would include information distribution where alliance members do have complete information about each other's preferences, the case when only one of allies (or a group of them) has a complete information on other partner's (or partners') valuations, and the case when nobody has the complete information on other allies' valuations[74]. This step includes a focus on a specific issue and the according goals and strategies accommodation between allies (we assume that a junction of these issues composes the total area of allies' common interest defined in the original accord[75]). The hypothesis

[72] Often shortened to *unk-unk*
[73] As this can be underlined as a crucial point of players' knowledge
[74] Simplifying the problem, we have narrowed down the analysis to two possible players with respective preferences, i.e. to two classes of players (practically, those might be groupings within alliances).
[75] The term 'original accord' does not refer to any specific time period, presumably at the point of alliance formation, as new issues that entail cooperation among allies can rise any moment during the existence of alliance. What it means is that all those issues become a part of the core agreement on common action, as concluded by allies in any form.

inferred from the previous discussion is that a lack of information in the game (whatever the deficient information distribution is) induces uncertainty and negatively affects the accommodation process. The less information available within the whole of alliance, the less is probability of reaching an accommodation point. It means that the situation in which all allies have imperfect information on each other's preferences is least likely to host an accommodation of goals and actions.

To prove these statements, we will adopt and slightly adjust Cédric Dupont's argumentation on incomplete information distribution within a broader sequential bargaining model he developed (Dupont, 1994, p. 162). The essence of our model appropriation is in reframing actors from simple random negotiators to allies, and in remodeling various parameters used by Dupont to fit our study-object more correspondingly (e.g. instead of prices, we will observe the case of accommodation costs).

In the case of complete information about the allies' valuations, we can first make the following proposition:

1. If the information distribution is even across the range of allies and they all reciprocally possess complete information about other allies' preferences, by using inverse rationality (backward induction) we conclude that the agreement is reached in the first stage of the alliance accommodation process, i.e. that no losses in gains from common actions are suffered.

   Schematic form of this proposition is:

   Period 1:

   A proposes x[76]

   B accepts any proposal with accommodation cost $c_x \leq x$

   Period 2:

   B proposes y[77]

   A accepts any proposal with accommodation cost $c_y \leq y$

   And its proof is:

[76] Certain action with cardinal value of projected accumulated common and individual gains x
[77] Certain action with cardinal value of projected accumulated common and individual gains y

Using inverse rationality, as suggested by Kuhn's Theorem, we begin with period 2: A accepts any proposal $c_y \leq y$, hence B proposes y. Therefore the payoffs of both allies are as follows:

$U_B(2) = \delta_B(b^{78} - y)$,

$U_A(2) = \delta_A(a^{79} - y)$

A step backward, we have B accepting any proposal $c_x \leq x$, hence A proposes x. The payoffs of both allies are then:

$U_A(1) = \delta_A(a - x)$

B accepts x if

$c_x \leq x$

and

$U_B(1) \geq U_B(2)$;

If not, the accommodation is reached in period 2 if

$U_A(1) \leq U_A(2)$

Therefore, the agreement is reached in period 1, as ally A has complete information on ally B's preferences and can incorporate them in its own proposal by suggesting x that implies $U_B(1) = U_B(2)$, as playing the second period is not rational, due to $U_A(1) \leq U_A(2)$ condition. This proof is based on a premise that both sides considerably value an accommodation.

Given the case of incomplete information on the side of one ally, we can make the following proposition:

2. Under one-sided incomplete information about other ally's preferences, some uncertainty is created, significantly affecting the accommodating process by decreasing odds for common action. Suppose uncertainty exists about ally B's preferences and ally A's preferences are common knowledge.

   Proving this proposition is directly inferred from the previous proof, as in this case ally A cannot suggest x that covers estimations of ally B's utility function. It means odds for suggesting x that implies

   $U_B(1) < U_B(2)$

[78] Ally B's most preferred utility related to the issue of accommodation that is being negotiated
[79] Ally A's most preferred utility related to the issue of accommodation that is being negotiated

are considerably high[80]. Suggesting *x* with such deficiencies in turn leads to B's refusal of accommodation in period 1. However, the accommodation is reached in period 2, as ally A's preferences are known to ally B, as previously said. That is why B can suggest *y* that implies A's relative utility function as follows:

$U_A(1) = U_A(2)$

However, if period 1 finishes without a point of accommodation, an agreement in period 2 as demonstrated is conditioned by both allies' high valuation of common actions within the alliance and of the alliance itself, especially higher than unilateral approaches encouraged by obstacles in the accommodation process. Moreover, irrationalities like sense of superiority or excessive flatulence of particular allies hamper this accommodation as well.

Considering the case of both allies having incomplete information about each other's preferences, we make the final proposition:

3. Under two-sided incomplete information, inefficiency in the accommodation process (and the subsequent losses) is largest if compared to cases 1 and 2.

   Again, proving this final proposition is inferred from the second proposition's proof. Particularly, given the failure of period 1 as in case 2, reaching a point of agreement in period 2 will in case 3 be analogously difficult and left to random guessing and estimation attempts with uncertain effects, as the probability of B suggesting *y* with implications

   $U_A(2) < U_A(1)$

   is as probable as the opposite case in the previous proposition.

   Consequently, reaching the accommodation point is hard to achieve, due to allies' overall incapability of calculating other players' utility functions, as their preferences are not commonly known.

This short examination of situations of imperfect information distribution has confirmed our claim on the causal relation between the level of information exchange among allies and the level of cooperativeness within alliance, including the probability of common actions and positions being agreed.

[80] The exact level requires further case specifications

Nevertheless, two methodological remarks related to the quality of this analysis should be made at this point. Firstly and specifically, this argumentation did not involve factors of domestic constraints on players, which would significantly extend the possible range of behavior of both allies. This factor was abstracted due to our exclusive focus on information distribution factor (and generally on intra-alliance information flow) at this point. Secondly and more generally, this analysis is limited as not only information distribution and information flows are relevant in the process of intra-alliance goals, positions and actions accommodation. There is a vast set of determinants influencing this process, as already established previously[81]. However, analytical reasons of a meaningfully specific focus required such a simplification. In other words, this brief analysis was not an attempt of building a middle-range model of intra-alliance sequential bargaining.

3.1.4 Payoff consequences (definition and preference ordering)

Regarded as utility awarding outcomes[82] of a player's actions chosen, i.e. strategy played, payoffs are meant to render different strategies analytically manipulable and classifiable. Less important is whether payoffs are defined in cardinal[83] or ordinal[84] terms. In terms of our study, this aspect of game theory does not pose major challenges apart from a general requirement for payoffs to be cautiously weighed, and the actually chosen modes (cardinal or ordinal) and subsequent designations (numerical values or descriptive terms) of payoff consequences to be well-argued.

However, a few discussion points arise from the assumptions preceding the payoffs framing stage of game construction.

Firstly, as suggested by game theory, ally members are expected to choose their strategies according to a well-established and fixed preference order of possible outcomes previously defined. Nevertheless, defining a preference order is rather complicated for a

[81] See chapter II, section 2. 2
[82] Utility indices
[83] With numerical values to outcomes of certain actions assigned and scaled
[84] Framed in descriptive terms (possibly using even figures in descriptive, non-scalable ways), such as 'satisfactory' and 'unsatisfactory'

state involved with an alliance. As we have shown in the previous chapter[85], determinants influencing allies' behavior and decision-making are complex and intertwined, often hard to discern and estimate in relative terms[86]. That is why allies' individual preferences might shift in the course of game at a pace that is sometimes hard to grasp by using the standard apparatus of game theory. Therefore, payoffs are not givens at all, contrary to what most game theorists would suggest (Allan & Schmidt, 1994, p. 10). Our instruction leads us towards an approach presupposing a highly refined study of factors preceding any identification of payoffs needed for a game to be played, especially an attentive investigation of changing determinants. For instance, not during entire course of the game will allies' relative military capabilities remain in the same constellation, nor will the relative influence of extra-alliance actors be constant. Consider an example of the Bulgarian-Yugoslav alliance of 1947, where the overall valuation of allying by both sides dropped once the Soviet Union introduced its anti-Yugoslav discourse and additional pressures on its satellites to take a stance of enmity towards the Yugoslav communist regime, opened by first disaccords between Belgrade and Moscow on the notion of the Balkan Federation and the ensuing Resolution of the Communist Information Bureau condemning the Yugoslav Communist party and its policies (Stankovic, 1954, p. 353). In this particular case, identifying dynamics of Bulgaria's preference ordering concerning particular intra-alliance-related actions[87] takes more than just defining stages of the game and the startup preferences, as they shift throughout the game at several occasions.

Secondly, according to game theory, allies form their preferences independently from each other. However, there are a number of historical examples that can be analyzed from a standpoint of *interdependently* formed utility functions and the according preference orderings. Those games regularly include three factors of interdependent utility functions: superiority, benevolence and malevolence, showing how individual player's satisfaction is drawn from or defined in relation to other players' gains or losses (Nicholson, 1994, p. 79). Consider, for instance, an example of Austrian-Russian axis within the Holy League. This Austrian-Russian formal alliance, intended to provide a

[85] See chapter 2, section 2.2
[86] By choosing predominant factor(s) once all the determinants' levels of impact are individually identified and subsequently compared with each other
[87] E.g. to support or not to support the Yugoslav Communist Party at the Communist Information Bureau meeting

legal ground for mutual military support of the two empires, was bilaterally established in 1833, and this treaty was invoked during the Hungarian revolution of 1848/49, when Russian troops entered Moldavia, Transylvania and eventually Budapest itself to support the Austrian ally against Hungarians (Istvan Deak, 1979). Just illustrating the above thesis on interdependent utility functions, we briefly suggest a set of reasons that drove Russia to fight against the revolution in Austria. These have a twofold nature: one part related to Russia's primary and independent gains attainable by an intervention in Austria, and the other part, definable as contextual gains, related to the satisfaction Russia drew from Austria's individual gains and satisfaction, as outlined in Table 1.

Russia's *primary* gains	Russia's *contextual* gains (*drawn form Austria's independent primary gains*)
• Fulfilling the 1833 Treaty obligations • Helping Austria in order to prevent a spillover effect of the revolution on Russia • Keeping the balance of powers in the XIX century Europe	• Keeping Austria politically alive for the sake of the European Concert • Russian tsarist devotion to conservative values and monarchical brotherhood

Table 1 Russia's primary and contextual gains from its intervention in Hungary in 1849

The above example shows how the factor of *benevolence* (related to Austria) on the part of the Russian emperor Nicholas I influenced his vision of expected gains for his empire[88], changed the overall Russian preferences in this case and led to the Russian intervention in Hungary in 1849. More generally, it proves that allies can form their preferences interdependently, in relation to other allies' (dis)satisfaction. It consequently means that the presumption of game theory on independent preference formation has to be redefined in order to be constructively incorporated into our meta-model. Following the above logic, we suggest this presumption's redefinition (when applicable) by a

[88] By amalgamating Russia's primary and contextual gains in the game

respective inclusion of such reframing factors as superiority, malevolence and benevolence, as these factors may play a salient role in intra-alliance relations.

Thirdly, there is some flaw about the way game theory treats payoffs at the basic level of its application. Namely, according to game theory, allies are set to treat their expected payoffs the same as certain payoffs, i.e. their estimations are seen as risk-neutral. This assumption, however, may not hold in practice. For instance, a risk neutral ally should value equally certain military support from its allies[89] as 30% odds of military support and 70% odds that they are not going to fulfill their obligations. Let us see how this is rebutted by simply observing utility functions.

Let x be a payoff consequence of allies' military support to a state A. Let $U_A(x)$ then be a continuous and increasing function of x. The function $U_A(x)$ gives state A's level of satisfaction in fictional "utils" from receiving payoff consequence x, and is known as a *utility function*. If the certain payoff of 100 utils is preferred by A to the gamble, (due to risk aversion phenomenon) then A wants a utility function that satisfies:

$$U_A(25) > .30^{90}\ U_A(100) + .70^{91}\ U_A(0)$$

The left hand side of this relation represents the *certain payoff* for A, whereas the right hand side is the *expected utility* from this gamble. Using an exemplar concave function like $U(x) = \sqrt{x}$, we come to the following form of the above equation

$$\sqrt{100} > .30\sqrt{100} + .70\sqrt{0}\ , \Leftrightarrow 10 > 3$$

As demonstrated, the certain payoff's utility is higher for the state A ($U_A(x) = 10$) than the expected utility of the risky payoff ($U_A(x) = 3$). This simple proof, in turn, shows that holding payoffs rigid and risk-neutral does not correspond to everyday reality. It also shows how computing the factor of *risk aversion* when framing payoff consequences for allies' possible strategies seems to be highly justifiable, as leaving this factor aside may

[89] Being a collective defense case
[90] As the probability of allies helping out is 30%
[91] As the probability of allies failing to support is 70%

render impossible explaining why allies refrained from certain strategies, although they were indicated to be their most rational choice.

As an empirical proof in favor of our remark, we should seriously take into account a case of Italian alliance with Germany and Austria-Hungary that broke apart in 1915[92], after Italy realized that the probability of acquiring territories targeted in the Adriatic basin was significantly higher if Italy joined the Entente Cordiale, i.e. that the risk of remaining empty-handed was appreciably lesser if compared to the risk of expecting territorial-wise compromises with Austria-Hungary. Given Italy behaving as a risk-aversive ally, it eventually switched the side and joined the Entente.

This segment of our discussion has shown us possible new and broader views on payoff consequences definitions that we, if the prerequisite of logical and substantial coherence is fulfilled, intend to incorporate into the meta-model. Those are above all the notions of shifting payoffs examination, interdependent utility functions and the role of probability, uncertainty and risk in framing payoff consequences.

3. 2 Process of cooperation in terms of game theory

A part of the link with the notion of intra-alliance cooperation, its possibilities and perspectives is made in the game information section of this chapter. However, it is still needed to demonstrate the link between game theory as an analytical framework and the substantial process of cooperation within international politico-security alliances. This link will be made through an observation of the ontological core of alliances – allies' *joint decision-making*, done through a game of mutual accommodation of preferences.

The question to be asked before we continue with this portion of discussion is: Why do we need to regard the process of collective decision-making to be able to consequently apply game theory to the examination of intra-alliance relations, given that any specific game[93] can be applied even without such a strict condition of additional analysis? The answer is somewhat already contained in our so-far analysis. Namely, the

[92] For more on this case, see Renzi, W. A. (1968). Italy's Neutrality and Entrance into the Great War: A Re-examination. *The American Historical Review*, Vol. 73, No. 5, pp. 1414-1432

[93] Prisoner's dilemma, stag hunt, chicken game (dove-hawk), coordination game, deadlock etc.

first portion[94] of our theoretical input has adopted the stance that international alliances are an entity different from a simple aggregate of its members, and that the collective level of decision-making is made of a process of mutual interests, goals and actions accommodation, i.e. it represents the core of all intra-alliance processes. Disregarding that remark at this point would mean simply leaving the game theory portion of the input incomplete and unrelated to the reasonably suggested and well-argued structure of the meta-model built up so far. Moreover, it would mean arbitrarily brushing aside a major element of the intra-alliance relations' dynamic. We cannot simply focus on the bare process of intra-alliance interacting without asking how allies reach their decisions and what it means for the framework of game theory.

This section's discussion is underpinned by the portion of argumentation we used beforehand to examine whether international alliances can produce independent effects on the state of international politics and their actors, including the allies themselves. We will follow the same course we established then, by firstly looking at the standard rational choice theory explanation comparable with the neorealist assumptions on international politics. We will then move on to a criticism of this rather narrow view of intra-alliance accommodation by presenting arguments that treat alliances as a context of a truly genuine *collective* decision-making that creates effects distinguishable from mere individual states' preferences, strategies and actions with certain consequences.

3.2.1 Preference aggregation

The case for this approach has been built since the startups of rational choice theory. The statement of this case, related to our topic of intra-alliance cooperation, is simple:

- Each state has its independent and fixed preferences;
- Intra-alliance decision-making is reached through a process of preference aggregation that results in a decision on specific actions allies are to take jointly;
- This aggregation process presupposes mechanically adding each state's preferences to other allies' individual choices, which means that each decision

[94] Chapter 2

made within an alliance can, through a disaggregation process, be traced back to particular preferences of allies (Guillermo, 2001).

We will suggest that within most alliances decisions are being made through majority voting[95], defined as alliance choice rule *f* that produces an alliance-level preference order *R = f (P)* such that:

$$\forall x, y \in X,\ xRy \Leftrightarrow N(xP_i y) \geq N(yP_i x),$$

where *i* stands for individual alliance member, X for a set of strategy choices, N for a number of alliance members and $N(xP_i y)$ designates the number of states *i* favoring *x* over *y*. Herein, a decision made at the alliance-level is identified with the existence of a fully transitive preference order at the according level.

Defining a dependent variable called *alliance profile* will at this point help us complete the picture of preference aggregation game-model of intra-alliance decision-making.

Therefore, an alliance profile *P* is a set of restricted individual states' preference orderings P_i *(i* $= 1, ..., N-1, N)$. In the special case $N = 1$, the alliance collapses into a *unitary actor*[96] profile.

Given this profile *P* and its aggregated order *R = f (P)*, an alliance-level decision exists *iff* there is a full rational preference order *R'* such that:

$$\forall x, y \in X, xR'y \Leftrightarrow xRy$$

Then we call *R'* the alliance-level decision made of aggregated preferences[97][98]. Closing this argument, we have shown how the mechanical process of preference aggregation results in a decision taken at the level of alliance. This decision, by definition is made of individual states' choices, and does not reflect any specific transformative power and

[95] Laws of international anarchy and sheer interests play a role as much as democratic principles. Both at this point of the discussion underpin this presumption.
[96] Basically corresponding to the assumption of a single ally-leader
[97] Given this discussion's general purposes, we will not deal with the Condorcet voting paradox and the assumption of single-peaked preferences (Arrow, 1963, p. 77).
[98] Adapted from Allan & Schmidt, 1994, pp. 55-65

influence of such class of institutions as international alliances. This portion of the discussion is therefore an extension of the neorealist denial of international alliances' independent impact on the state of international politics and its actors. However, the next section will show how intra-alliance cooperation involves allies in a genuine process of goals and preferences accommodation rather than simple aggregation, therefore building a direct link with the first strand of our theoretical input, where a compromising idea of international alliances as institutional contexts for policy accommodation with separate and distinguishable effects is made[99].

3.2.2 Intra-alliance preference accommodation

A harsh critique of the approach sketched above was introduced by scholars (Hodgson, 1967; Gauthier, 1975; Regan, 1980; Bacharach, 1993, 1999; Hollis, 1998) to show how producing decisions at the collective level takes more than arbitrarily choosing an aggregation principle[100] (like the majority voting rule) and simply mechanically adding individual allies' preferences one to another.

Simply put, the above assumptions on 'alliance A prefers *x* to *y*' are, contrary to the preference aggregation model, not the same as 'each member of alliance A prefers *x* to *y*'. This remark bluntly explains the main course of the critic of the traditional rational choice approach and calls for theoretical refinements.

The root of the controversy is traditional rational choice theory's assumption that only individuals, and herein states, can face decision-making problems, not the supposedly derived entities as groups, herein international politico-security alliances are. Once the decision-making problem is faced and solved at the level of alliance member, it only remains to add such particular decisions together and receive the overall decision. This notion is perfectly in the line with the neorealist state-centrist credo, where no room for authentic contribution of alliances is left.

[99] See chapter 2, section 2.1.4

[100] Tuomela (1995, pp. 185-191) defines this aggregation principle as a specific transformation *procedure*, i.e. as a method or a formula that transforms individual preferences into collective ones, naming it formally "transformation function". He, however, does not specify this procedure in any detail.

However, Arrow's impossibility theorem[101] has generally shown that the simple aggregation model leads to a dead-end (Arrow, 1963), as it cannot fulfill the basic logical conditions posed before it. His and the works that followed (Sugden, 2000) confidently argued that the rational choice theory should regard collective entities, in our case international alliances, to be decision-making agents, which, in turn, leads to a broader possibility of a specific application of game theory in the analysis of their interactions. It consequently means the scholars suggested that, with regards to our case, alliances can have their genuine preferences irreducible to preferences of individual allies' simply taken together. The reasoning proposed implies that the preferences that alliance members have if taken collectively are *distinct from*, but *on a par with*, the preferences that guide their individual choices.

In accordance with the above remarks, we broadly suggest the *team-directed reasoning model* (Sugden, 2000, p. 191; Bacharach, 2006, pp. 120-155) of derivation of collective preferences as an explanation of how decisions are reached at the alliance-level. Specifically, team-directed reasoning is a model of reasoning that players use when they take themselves to be acting as members of teams. Then, according to the model, they do not set their preferences in line with their individual likings, but rather think of what should the preference order look like given the team's goals. This concept obviously overcomes the mechanistic preference aggregation logic, as previously depicted. Herein, the team-directed reasoning model refers to alliance members and alliances as their group and institutional context. Furthermore, this model implies that an alliance exists to the

[101] The theorem, also known as "Arrows paradox" basically states that no voting system that suggests aggregating individual preferences can possibly meet a putative set of reasonable criteria if three or more options are given to choose from. This preferential voting system is labeled social welfare function, and the pertaining criteria, as defined by Arrow are:

1. **Universal domain constraint (U)**: the social welfare function should account for all preferences among all voters to derive a unique and complete ranking of collective choices. Thus, the voting system must account for all individual preferences, it must do so in a manner that results in a complete ranking of preferences for the chosen collectivity, and it must deterministically provide the same ranking each time voters' preferences are presented the same way;
2. **Absence of a dictator (D)**: excludes the possibility of a dictator, the social welfare function cannot simply follow the preferences of a single voter;
3. **Independence of irrelevant alternatives (IIA)**: adding an alternative to a set of alternatives should not affect the collective ranking of the previous options;
4. **Strong Pareto condition (P)**: if each voter (herein alliance member) prefers an alternative, then the alternative is collectively preferred;
5. **Transitivity (R)**: there must be a consistent collective ranking of alternatives (Arrow, 1963; Allan & Schmidt, 1994)

extent that its members take themselves as a part of it, which takes us back to our inferences about international alliances drawn from neorealist thinking, i.e. to a notion that alliances exist only as long as allies' interests are converging in an issue-area. However, our model has upgraded this view by incorporating arguments that emphasize allies' interests to be engaged in a process of intra-alliance cooperation[102]. Similarly, this model does not imply any neorealist normative stance which makes it possibly compatible with our meta-model.

Particularly, if a set of individuals A_1 ,..., A_{n-1}, A_n is given, we can frame a following game: For each A_i there is a set Si of alternative strategies, from which A_i must choose one and only one. Consequently, every possible array of chosen strategies (one chosen by each individual), has an outcome. Furthermore, we define an *alliance-conditioned utility function* $a(.)$ which assigns to each x outcome a utility index $a(x)$ that we decide to call *alliance-conditioned utility*. We give an example in table 2 that represents this game as a coordination problem. Two allies lobbying two different states within an international organization is a setting of the game[103]. Allies have to choose between states P and Q, but there is only one scale of preference on which outcomes come to be ranked: *the scale of alliance-conditioned preferences*.

/	P	Q
Q	10, 10	0, 0
P	0, 0	11, 11

Table 2 Alliance-conditioned coordination problem

Now, let G be any game form and let $a(x)$ be an alliance-conditioned utility function. Suppose there is a set of strategies $(s_1^*,...,s_{n-1}^*, s_n^*)$, where each s_i^* is an element of S_i and where, given $a(x)$, the alliance-conditioned utility derived by this combination is strictly greater than utility created by any other combination. In that case, each A_i reasons along the lines of the alliance-conditioned problem with respect to G and $a(.)$ if it goes for s_i^* in the light of the fact that the set $(s_1^*,...,s_{n-1}^*,s_n^*)$ uniquely maximizes alliance-conditioned utility. Important to notice is that allies do not set their

[102] See chapter 2, section 2.1.4
[103] For details, see footnote 51

alternatives according to their individual preferences, but according to the preferences derived from alliance-level objectives[104].

Defining this game as a coordination problem[105], we have defined and generally depicted an alternative view on collective preferences formation. It is namely alternative to the traditional preference aggregation model and possibly more adequate in explaining the reality of intra-alliance relations, as it sets members' alliance-level preferences as distinct, but compatible with the individual ones.

However, invoking the balanced conclusions on the nature of the process of intra-alliance cooperation we drawn in chapter 2, we must shortly discuss the issue of adequateness of these two alliance-level preferences formation models[106]. As we have already stated, the preference aggregation model is logically slightly more in line with the neorealist assumptions on international cooperation, whereas the model of team-directed reasoning recognizes the transformative and independent role of international alliances as international institutions, mainly assigned by Liberal Institutionalists. Which model is in a specific case going to be applied predominantly depends on the type and state of alliance in question (i.e. the range of issues it covers, temporal orientation, patterns of intra-alliance relations[107]). Once that state of relations is broadly defined, a researcher can opt for a specific preference formation model. However, we would at this point recommend a rather regular deployment of the team-directed reasoning model, as it is far more comprehensive and could, with smaller adjustments and methodological remarks, be applied to alliances with highly 'selfish' alliance members as well.

With this observation, we close the portion of our discussion that was meant to show how alliance-level decisions are being taken in terms of game theory's analytical apparatus within a broader process of intra-alliance cooperation.

[104] Adapted from: Sugden, R. (2000). Team Preferences. *Economics and Philosophy*, 16, pp. 175-204
[105] Arbitrarily chosen for the sake of clarity
[106] Namely, preference aggregation and team-directed reasoning model
[107] For more on that, see chapter 2, section 2.1.4

4. APPLICATION CAPACITY OF GAME THEORY IN TERMS OF INTRA-ALLIANCE RELATIONS: COMPLETING THE META-MODEL

This chapter will try to offer an answer to the question of game theory's applicability to a study of intra-alliance relations from a perspective based on the predominantly theoretical argumentation presented in the course of our meta-modeling enterprise. However, an empirical testing of the meta-model will be done in the next chapter. At this point, a rather methodological assessment will be performed.

Assessing the application capacity of our meta-model will be done by firstly completing the properties of the meta-model, and, secondly, by setting out the criteria of its *descriptive*, *explanatory* and *predictive* (or prognostic) value.

4.1 Meta-model of analysis of intra-alliance relations: the completion

The meta-model will be completed by merging input1 and input2 constructed in the course of the paper. Following the criteria of logical and substantial coherence, we will outline the 'compatibility strings', i.e. the connections that render the whole of the meta-model internally coherent.

Condensing two sets of theories will be done by attaching their properties at the compatibility points provided by the ontology of meta-model projected in chapter 1.

Once more, the meta-model's ontology is projected to be twofold, encompassing:

1. components of international politico-security alliances within the game theory framework[108] (*C-dimension*);
2. determining factors of international politico-security alliances in terms of game theory elements (*DF-dimension*);

This ontology outline entails the very ontology of the meta-model. As claimed throughout the paper, this ontology contains both structural (static) and processual (dynamic) features at various levels of its substance.

The C-dimension therefore acquires the following design:

[108] The methodological sequence is as follows: international alliances' empirical features abstracted to the level of independent variables – theories of international alliances as first class intervening variables – game theory as a second class intervening variable – meta-model as dependent variable

1. *Alliance members* – systemic actors as *players* in game theory (see chapter 2, sections 2.1.1, 2.1.2 and 2.1.4, and chapter 3, section 3.1.1);
2. Fundamental *alliance accord – legal game regulator* that sets out a portion of rules of any game (see chapter 2, sections 2.1.4 and 2.2), as derived from allies' mutual expectations – it is a part of the *information set* allies posses in a game (see chapter 2, sections 2.1.1 and 2.1.4, and chapter 3, section 3.1.3) ;
3. Power flow and capabilities distribution within alliance – *pragmatic*[109] *game regulator*, de facto forming rules of any intra-alliance game (see chapter 2, sections 2.1.1, 2.1.2, and 2.1.3, and chapter 3, sections 3.1.2 and 3.1.3)
4. Allies' common and compatible *interests* (individual, collective, security or autonomy-driven) – directly related to *payoff consequences* and *preferences over payoffs* in game theory (chapter 2, section 2.1.1 and 2.1.2 specifically, and chapter 3, sections 3.1.4 and 3.2);
5. Intra-alliance *cooperation* as coordination and *mutual accommodation* – conclusions and insights deployable in terms of game theory when defining *strategies* and *tactics*, and when demonstrating the process of intra-alliance *preference accommodation* (chapter 2, sections 2.1.3 and 2.1. 4, and chapter 3, sections 3.1.2 and 3.2).

The DF-dimension of the meta-model has been previously built up to define a set of determinants that decisively influence the behavior of alliance members[110]. These determinants, particularly, shape the crucial logical link of any game: the one between payoff consequences formation and formation of preferences over payoffs (their rankings), while both in turn explain strategy choices of alliance members. Moreover, each of these determinants has its own specific mode of manifesting, which is briefly summed up[111] in the following draft:

1. *Intra-alliance stimuli:*

[109] 'pragmatic' in terms of its systemic or meta-model internal role, not in the sense of allies behavioral qualities (as related to their capabilities)
[110] However, some room for unpredictable occurrences and spontaneous events should be left, as this paper does not claim to follow the path of scientific determinism
[111] Being preceded by a more thorough depiction in the previous course of the paper

- Allies' varying *interests* and motivation to ally with other actors – determine the level of competitiveness and cooperativeness of a game[112]
- Allies' particular definitions of *national interest* – colliding national interest may turn, for instance, coordination into a prisoner's dilemma game, i.e. hinder intra-alliance cooperation, or even disrupt an alliance[113].
- Individual member's internal politics, socio-economic, demographic and military conditions, political and regime similarity of allies – this vast set of conditions can explain why states sometimes value indirect, or simply lesser gains, more than direct, or larger ones[114] ;
- Relative significance of issues with converging/diverging interests among allies[115], resulting parallelogram of allies' diverse converging/diverging interests – affects the very nature and cooperativeness level of a game;
- Allies' relative capabilities – defines the interests of allies, influences their ambitions in terms of preferences formation, defining expected payoffs and their ordering;
- Allies' strategies of intra-alliance behavior – influence the mid and long-term strategic choices of allies (e.g. if the strategic logic of *quid pro quo*, or tit-for-tat is deployed);
- Intensity, extensity and density of information flow among allies – dominantly influences the state of information distribution in a game, entailing the expected payoffs and subsequent strategies;

[112] Think of the Soviet threat Western Europe was facing during the Cold War and how it led to tight transatlantic relations, and, then, compare it with the reappearing tensions once the Soviet threat has been removed. We could frame a post-Cold War game among allies as being considerably less cooperative than the Cold War one.

[113] Consider the example of the Habsburg-Italian relations at the very beginning of WWI, and how due to the colliding definitions of national interests (both states wanted the Dalmatian sub-region and the eastern Adriatic cost) a long-term alliance between these two states could not be preserved, which at the end led to the Italian withdrawal from the Central Power axis and to Italy actually fighting Austria-Hungary (their former ally) quite ferociously

[114] Ideological proximity, shared or tightly related national identities (think of how it brought Serbia close to Russia in wars against Turkey at the beginning of XIX century, although Russian army often failed to substantially support the Serbian troops (Trivanovitch, 1931).

[115] Recall how the increase of relative importance of transatlantic allies' differences on the level of threat Saddam Hussein's posed regionally and globally weakened that transatlantic alliance over time.

- Allies' individual and common historical and security legacy – by affecting the level of trust and nature expectations that allies share, it indirectly affects their specific moves as well.

2. *Extra-alliance (environment's) stimuli* (number of extra-alliance actors and fundamental and situational set of military, political and economic challenges, risks and threats) – all can, first, (directly or indirectly) influence the level of cooperativeness of a game, by giving incentives or disincentives for cooperation[116], and, second, they can crucially change the nature of a game (e.g. inclusion of an additional, previously extra-systemic actor as a player can make a shift from, for instance, dictator to a screening game[117]).

[116] If an ally finds collaborating in certain areas with actors outside of alliance individually more beneficial, thus neglecting the intra-alliance commitments.
[117] As the additional player can undermine the role of dictator from the original game

5. APPLYING THE META-MODEL TO A CASE STUDY: TRANSPACIFIC RELATIONS IN LIGHT OF THE NORTH KOREAN NUCLEAR CHALLENGE

"*The only thing worse than fighting a war with allies is fighting a war without them"*
(*Winston Churchill*)

The goal of this chapter is to test the applicability and relevance of the meta-model previously constructed. This will be done by examining the case of relations within what is denoted here as the Transpacific Alliance, involving the US, South Korea and Japan. Their relations will be analyzed in light of North Korea's challenge to the authority of the Nonproliferation Regime and security of the Alliance. Specific focus is placed on two nuclear crises on the Korean Peninsula induced by North Korea's announcements to withdraw from the Nuclear Non-Proliferation Treaty (1993-1994 and 2002-2007), and in the case of the second crisis – by its actual withdrawal.

The fact that after the collapse of communism throughout the world, North Korea remained the most immediate and crucial threat to the security of South Korea, Japan and US forces in these two countries makes its impact on these states' intra-alliance politics large both theoretically and practically. In spite of the regional presence of China and Russia, North Korea remains that most imposing reason for the preservation of the Alliance, at least in its state of high alert and deterring military presence of the United States.

Precisely due to this fact, the final section of this study looks into how North Korea's behavior, strategic and tactical decisions drove the US-South Korean-Japanese intra-alliance relations. It examines crucial phases of intra-alliance relations throughout the two crises, tries to determine tendencies in allies' decision-making on how to respond to the nuclear challenge and offers cautious projections on future developments of those relations with respect to the North Korean nuclear issue and all-Korean unification as a part of the riddle.

5.1 Introducing the Game

Before commencing the description of this intra-alliance game, arguments in favor of North Korea's relevance as a security threat to South Korea, Japan and the US will be given. This argument will prove North Korea's behavior should be studied as a determinant of relations within the Transpacific Alliance.

Firstly, North Korean regime has been repeatedly and explicitly, both publicly and off the record threatening the security of South Korea, Japan and US forces in these two countries. Recalling the "sea of fire" threat made by North Korea during the inter-Korean Panmunjom talks in March 1994, it should be emphasized that its regime made one thing clear to the allies: North Korea is ideologically, economically and politically different, it is there to stay, and if felt immediately jeopardized, it is ready to use force (Wit, Poneman & Gallucci, 2004, p. 149). Moreover, if any unification is to take place, it will be done under the terms of North Korea. This rigidity calculatedly suggested an implicit threat of North Korea's invasion of the South (Downs, 1999, p. 192). Finally, one should not forget the North's blitzkrieg of 1950 and the inception of its "Fatherland Liberation War", its sudden nature and overall consequences.

Furthermore, North Korean military and destructive capabilities should be coupled with its unclear offensive intentions and its continuing sense of insecurity while surrounded by the US proxies. It is a compelling fact that the North has a clear preponderance in conventional forces and manpower in the region. North Korean active troops are fourth largest in the world with over 1.2 million personnel Compared to the US 37.000 troops in the region, South Korea's 600.000 and Japan's 235.000 troops, it becomes clear that North Korea has about 1.5:1 advantage in manpower. In addition to this statistic, DPRK has almost 7.5 million people under paramilitary reserves, meaning that some 40 percent of the population serves in military or paramilitary reserves. Although its military planning and intentions remain uncertain, the way North Korea has been deploying its forces along the Demilitarized Zone indicates they could transform into invasion units supporting a total assault on Seoul, especially given DPRK's massive artillery batteries only 24 miles away from the South Korean capital. This assault could be buttressed by 65.000-man strong Special Forces deployable for the purpose of

destroying South Korea's nuclear reactors, being the economic key to its security and an environmental hazard alike. North Korea also has a 2:1 advantage in tanks and airframes over the allies, which only adds to a complete picture of its conventional predominance over the Alliance. However, old equipment, old-fashioned training and malnutrition of the North's military undermine its potency vis-à-vis the US, South Korea and Japan (Kyoung-Soo, 2004, p. 22; Sanford & Scobell, 2007, pp. 22-23; Wit, Poneman, Gallucci, 2004, p. 180).

Moreover, US Intelligence made several different estimations on North's nuclear capabilities, pondering whether it had already produced a number of bombs and what its aggregate capacity to produce them actually is. In November, 1993, the Intelligence community assessed that there was a "better than even chance" that North Korea had already produced one or two bombs (Wit et al., 2004, p. 128). On the other hand, more general and mid-term estimates said there was a prospect of North Korea fabricating about five bombs back in 1994 (p. 103), up to six to eight bombs in 2006. However, on February 10, 2005, North Korea announced it had actually manufactured nuclear weapons, which was to confirm its active and deployable nuclear potentials. This proclamation had been widely doubted and interpreted as a bargaining tactic before DPRK conducted an underground nuclear test in the vicinity of P'unggye (northeast of the country) on October 9, 2006, which was confirmed by the US Office of Nuclear Intelligence, too. While all the estimates indicated those were plutonium-based devices, it remained unclear whether they are missile capable. It has been, however, suggested that North Korea has not yet acquired all the necessary delivery technology. Moreover, those devices yielded an explosion four times weaker than expected (less than 1 kiloton as opposed to planned 4 kilotons), which leads to a conclusion that its nuclear weapons are at this stage still rudimentary, not posing an immediate threat (Hecker, 2006). Finally, nobody can say for sure if and how many nuclear weapons North Korea possesses at this stage and what their destruction capability is. Nevertheless, precisely this ambiguity and great deal of uncertainty leaves room to military and political speculations and opposing diplomatic strategies suggested within the Transpacific Alliance. It has also remained the prime diplomatic strategy of North Korea since 1992.

Along with its conventional military superiority and increasingly threatening nuclear weapons program, DPRK has a long history of ballistic missile development, too. This is particularly important as these missiles appear as delivery systems for its nuclear weapons. At this stage, North Korea possesses approximately 800 ballistic missiles of various generations. It is known to possess around 300 to 500 Hwasong-5/6 short-range (tactical) missiles based on the SCUD technology, with a range of up to 500 km. Concerning medium-range missiles, DPRK has up to 200 1.300-km-range Nodongs, enabling it to ballistically embrace the entirety of both the Korean Peninsula and Japan, including the two capitals, Seoul and Tokyo, military bases in Okinawa and Guam, power plants, major military and civilian objects. Finally, North Korea holds long-range Taepodong missiles capable of potentially reaching even the territory of the US. However, this threat appears to be only a hypothetical one, as its Taepodong-1 missile with a range of 1.500 to 2.500 km (still too short to reach the US territory) was unsuccessfully tested in August 1998, failing to enter the orbit due to a lack of power. Nevertheless, this test did stun the world, showing that North Korea managed to acquire all the key technologies needed to develop inter-continental ballistic missiles (ICBMs) capable of reaching even the US soil. This technology comprised multiple-stage separation, stable guidance systems of multiple-stage rockets, multiple fuel systems etc. Particularly alarming is the fact that DPRK has been already developing a type of ICBMs, Taepodong-2, with a range of up to 6.700 km. Deploying Taepodong-2 is however, projected to be technically possible only after 2012. Its current status is estimated to be at the level of a prototype testing (Yun Duk-Min in Kyoung-Soo, 2004, pp. 121-137).

Given the capabilities ratio presented above, the uncertain, yet burgeoning nature of the North Korean nuclear program and its history of military offensives on the southern neighbor, it is fully justified to analyze DPRK's behavior as a determinant of the state and dynamics of relations among the US, South Korea and Japan.

5.2 Allies playing games – sequential bargaining within the Transpacific Alliance

In this section, an overview of relations within the Transpacific Alliance during two nuclear crises will be given and explained using the tools of sequential bargaining games. Once the two games are completed, cautious predictions on future responses of the Alliance to potential challenges posed by North Korea will be given. This prediction will be based upon a computation of possible alliance-level preferences on a common stance to be taken vis-à-vis DPRK. These preferences will, in turn, be derived from past tendencies in intra-alliance politics driven by the, broadly speaking, North Korean issue.

5.2.1 First Korean crisis: March 1993- October 1994

The crisis opened on March 12, 1993 by North Korea's announcement of its intention to withdraw from the Nuclear Non-Proliferation Treaty did not mean just a crisis in relations between the United States and it allies and North Korea, or even between the IAEA and North Korea. This announcement also induced an at times serious crisis in relations among the transpacific allies: the United States, South Korea and Japan. Putting it bluntly, this intra-alliance crisis was due to a simple fact: although close allies in Northeast Asia, and even though a large portion of homeland security of South Korea and Japan depends on the military presence of the United States, the US, South Korea and Japan cannot always have identical preferences on all issues, including North Korea. This comes from a variety of factors; domestic politics and political opportunism, diverging security priorities, strategies and perceptions being just some of those. As no country or leadership is immune to cognitive dissonances and political entrenching, it is no wonder that close intra-alliance cooperation may lead to occasional disagreements. However, as emphasized earlier while constructing the meta-model, shared goals lead to common alliance actions in the long run, even though short term and tactical preferences may differ. Game theory enlightens this situation as a "myopic equilibrium" (Brams, 2001) and the meta-model fully counts on this course of explanation, as stated earlier. If these short term differences are not accommodated, alliances may dissolve in the absence

of unifying forces that make allies overcome their internal disputes. This unifying force did not miss in the case of the pressing security threat posed by North Korea.

North Korea's announcement caught the Alliance unprepared, bringing intra-alliance diplomatic contacts to a state of alert. The allies had ninety days to devise a response strategy before the North Korean withdrawal became effective. However, preventing abrupt American reactions soon became a top diplomatic priority for Seoul, as it feared rising tensions might lead to a full-fledged armed conflict. This position was implicitly communicated to the US through a South Korean push to terminate a major military exercise conducted jointly by the two allies – Team Spirit. This was finally done on March 19, when Seoul announced the conclusion of Team Spirit.

Therefore, the stage of intra-alliance bargaining is framed as an alternating offer game. Moving first, South Korea offers one initial way of crisis management – terminating Team Spirit. The US is to respond, moving second. Possible *physical outcomes* of the of this bargaining are vectors

$$t_{TS} = (t_C, t_{\bar{C}}, t_{RP}, t_{RE}).$$

The set of physical outcomes on the Team Spirit proposal (t_{TS}), thus includes canceling the exercise (t_C), not canceling and organizing it as planned ($t_{\bar{C}}$), rescheduling it as postponing it (t_{RP}) and rescheduling it for an earlier date (t_{RE}). As South Korea insisted on fully communicating the importance of canceling Team Spirit, the US has the information on South Korea's preference over outcome, knowing it prefers canceling it and values it the most. Moreover, both the US and South Korea knows the order of South Korea's preferences over outcomes:

$$t_C > t_{RP} > t_{\bar{C}} > t_{RE}.$$

Derived from physical outcomes are utility outcomes for South Korea as follows: $u_{SK_{C,\bar{C},RE,RP}} = U_{SK_{C,\bar{C},RE,RP}}(t_{C,\bar{C},RE,RP})$, and for the US: $u_{US_{C,\bar{C},RE,RP}} = U_{US_{C,\bar{C},RE,RP}}(t_{C,\bar{C},RE,RP})$. As for the US $u_{US} \geq u_{SK}$, it accepts the offer in the first period, and Team Spirit is canceled, giving each player its preferred payoffs.

Given allies' shared view that diplomatic solutions should be pursued, a series of coordinated attempts to engage North Koreans in negotiations soon started. However,

somewhat ambiguous position of the United States was encapsulated by the Assistant Secretary Clark in his conversation with the South Korean ambassador to the US:" We do not want to take any steps that would be seen as rewarding their behavior". This essentially meant 'yes' to negotiations, but 'no' to substantial concessions. Seoul and Tokyo, on the other hand, kept warning Washington that Pyongyang might view potential UN sanctions as an act of war and react violently (Wit et al., 2004, p. 32). This fear was buttressed by the US Intelligence, too, as it warned the administration of the same risk later that spring (Oberdorfer, 1997, pp. 259-60). This, in turn, limited prospective coercive measures available to the US who led the Alliance in its common dealings with the North.

These individual preferences resulted in an alliance-level preference to pursue a strategy of "gradual escalation" that would seek to build a broad international coalition and provide international support to the goals of the Alliance, thus increasing pressure on North Korea. As put by the Japanese, this strategy aimed a bit oxymoronically at being "firm but flexible" (Wit et al., 2004, p. 33). In late March 1993 in New York, all Allies agreed this strategy included a gradual escalation of sanctions, too, should North Korean concessions lack after a period pressures with international legitimacy.

At this point, having found about allies' preferences on coercion and sanctions, the US suggested the gradual escalation strategy, x_{GE} following the preference accommodation procedure of intra-alliance decision making. The strategy brought the alliance-level utility evenly distributed across the Alliance, and defined as:

$$a(x) = f(p_1, p_2, q_{NK}).$$

Herein, $a(x)$ is the alliance-level utility function, p the probability that the outcome x_{GE} will be entirely satisfying to the other allies and q is the probability that it will compellingly affect North Korea. As for the other two allies the utility of refusing the US alliance-geared gradual escalation strategy was lower than the utility of accepting it, South Korea and Japan accepted it, not taking this point of game to another stage, and avoiding the disagreement costs, i.e. the conflict payoffs c_{SK}, c_{Jap}. Thus, the gradual escalation strategy was adopted. This alliance-conditioned utility function, gained

through an intra-alliance preference accommodation procedure, is calculated for each player individually, following the formula:

$$a(x) = p_1 p_2 q_{NK}.$$

North Koreans, on the other hand, responded to this growing pressure by an international campaign of their own, announcing their conditions to return to negotiating table: permanent cancellation of Team Spirit, opening of US "nuclear military bases" in South Korea to inspection, a guarantee that the US would not launch a nuclear attack on the North, and restoration of IAEA impartiality. This announcement was another input to intra-alliance talks on how to respond to DPRK's actions, as it set the negotiating position of the North. Furthermore, North Korea's First Vice Foreign Minister Kang used a press conference to warn the Allies that referring the issue of its nuclear program to the Security Council could cause "grave consequences", thus putting the US, South Korea and Japan under additional pressure (Sigal, 1998, pp. 46-47).

Meeting in New York City on March 22, Allies' representatives to the UN agreed they could not determine whether DPRK was using its withdrawal from the NPT as a high-risk bargaining tactic or as a way of protecting its nuclear weapons program. They also agreed they would consequently need a response strategy that would be equally compatible with both options. Although dilemmas on the modality of talks with North Korea remained, a consensus that they were imperative was reached. After a series of signaling games and a somewhat surprising cooperativeness by North Korea, the US was on the threshold of negotiations with it by May 1993 (Wit et al., 2004, p. 39).

The allies' decision to pursue negotiations was a part of the gradual escalation strategy, so its alliance-conditioned utility mirrored the one adopting the gradual escalation strategy in general.

$$a_N(x) = f(p_1, p_2, q_{NK}),$$

with $a_N(x)$ being alliance-conditioned utility of negotiations with North Korea, p_1 and p_2 the probabilities that pursuing negotiations would be acceptable for the other two allies and q_{NK} the probability that negotiations would change North Korea's behavior. For each ally, this alliance-conditioned utility is calculated as following:

$$a_N(x) = p_1 p_2 q_{NK}.$$

A major part of the mosaic was an American effort to secure an inter-Korean dialogue. This was to seal the American commitment to the interests of its ally, acknowledging its preoccupations as well. As the focal point of any crisis and its resolution likewise were US-DPRK negotiations, South Korea had to rely on its ally to pave the way for its talks with the northern neighbor. Therefore, South Koreans felt they had to move in tandem with Washington, fearing the attention might be entirely diverted to the US-DPRK talks, thus marginalizing their own efforts. Their nightmare was that North Korea could use talks with Seoul for propagandist purposes or to avoid international criticism if the US dialogue failed (Wit et al., 2004, p. 46). This paralyzing fear was intensified by North Korea's Kim Jong Il stating negotiations between the two Koreas would begin after discussions with Washington had been set up. This sent a clear message to the South: any inter-Korean dialogue is of secondary importance to the North and largely depends on the pace of negotiations with the US. Americans, on the other hand, clarified that North-South talks had to be part of a broader diplomatic strategy. The US sent a signal to its ally that Seoul should diplomatically "reengage the North" (p. 47). Subsequent South-North contacts resulted in North Korea's initiative for deputy prime ministers to meet, ranking this proposal diplomatically as an exchange of special envoys. Their task would be to arrange a summit meant to resolve differences between the two countries and agree on the implementation of the joint Denuclearization Declaration (p. 64). The summit was not arranged due to an inter-Korean propaganda war and equivocal tactics deployed by South Koreans. Still, the South decided to remain flexible on the special envoys exchange proposal, as it believed it was the best way of supporting the New York efforts by the US (p. 65). Contacts between the US and North Korean diplomats through the New York channel, on the other hand, led to more ambitious talks in June 1993.

As the course of the game suggests, South Korea demanded from the US to link North-South dialogue to the US-North Korean negotiations. The set of physical outcomes was therefore $t = (t_a, t_{\bar{a}})$, where t_a meant the US accepting this proposal and $t_{\bar{a}}$ not accepting it. Again, the US responded positively. However, it is known that the US kept maintaining this link in deference to its ally, whereas it occasionally preferred to break

the link and not condition its talks with North Korea with inter-Korean dialogue. Applying backward induction, the US could come to the following conclusion.

Rejecting South Korea's offer bringing to all the three allies gains that corresponded to the utility vector $v = (v_1, v_2, v_3)$, meant proposing solution that would produce utilities $w = (w_1, w_2, w_3)$, such that v and $w \in \delta F$, with F being the feasible set of utilities. No conflict payoff would be received at this point, as the US would suggest a solution with the w utility vector. Only in case of South Korea's rejection of that proposal, the conflict payoffs, c_1, c_2 and c_3 would be received. The imposing question is why the US decided not to go to the stage 2 of bargaining and accept South Korea's proposal. Through backward induction, and being aware of South Korea's insistence to not to be sidelined in any negotiations, the US knew that $w_1 < v_1, w_2 < v_2$ and $w_3 < v_3$. As South Korea would fiercely reject any third option, the third stage of bargaining would not be possible either, so at the second stage, all the allies would receive conflict payoffs c_1, c_2 and c_3. As these conflict payoffs would designate a phase of serious discord within the alliance, and would be significantly lower, the US decided not to go there, even if conflicting with its ally in the short run would bring a faster solution in talks with North Korea in the mid-run. However, as these benefits were far from sure, the US decided to accept South Korea's proposal at the first stage, which brought utilities $v = (v_1, v_2, v_3)$. Here, South Korea decided to stick to its first offer, as

$$(1-p)v_{SK} + pc_{SK} \geq w_{SK},$$

with p being the probability the US would accept South Korea's offer. As South Korea knew the US would probably not continue with the bargaining, it decided to stick to its offer. Herein, South Korea was taking only its individual interest into consideration, without paying much attention to the alliance-level utility, so this utility consideration, if there was any, can be disregarded. The US, on the other hand, defined its utility vector taking into consideration what the alliance-level preference would be, i.e. what might be best for the alliance. It meant:

$$v_{US} = f(m_{SK}, m_{Jao}, q_{NK}),$$

where m_{SK} and m_{Jap} designates the probabilities that South Korea and Japan would find the solution bringing these utilities favorable to them, and q_{NK} designates the probability that this outcome would compellingly affect North Korea. Note that q_{NK} was in this case particularly low, significantly decreasing America's utility and gain from the decision to link its negotiations with North Korea to inter-Korean dialogue. It was low because North Korea strongly favored talking to the US alone. It should be emphasized that the utility vector w defined by the US and that could have been proposed at the stage two of this part of the game, was not an alliance-conditioned utility vector, as it reflected only the Americans' occasional preferences to talk to North Korea only without a prior condition of inter-Korean talks. However, the US did generally prefer North Korea talking to the South, as it was well aware that no viable solution could be found unless it belonged to everybody, South and North Korea alike

Following days of negotiations, the US and North Korea agreed on the so-called June 11 statement, laying out principles that would lead to DPRK's unilateral suspension of withdrawal from the NPT. Americans assured North Koreans they would not use or threat to use force (including nuclear weapons), respecting North Korean sovereignty and not interfering in North's affairs[118].

Although cautiously approving the statement, Japanese and South Koreans said that if DPRK believed they were completely satisfied, it was not true. Moreover, soon after, pressed by public criticism, both US allies would start expressing a negative opinion on the June 11 statement, condemning what they saw as a US-DPRK bypass of inter-Korean dialogue as an integral part of any solution with the North. The Americans thought this feeling was irrational, as they believed they did everything possible to keep their both Northeast Asian allies engaged, including daily meetings between Washington's main negotiator, Robert Gallucci, and Seoul's UN ambassador, Yu Chong Ha, as well as regular contacts between the State Department and the Washington embassy of the South. This feeling of the US negotiators and decision-makers was conveyed to Seoul. Additional warnings by the South Korean President Kim to the US

[118] *US-North Korean Joint Statement*. June 1993. Geneva

not to make any substantial concessions to the North clearly depicted a potential rift between the allies (Wit et al., 2004, pp. 66-67).

Herein, another South Korean proposal to the US is identified. Out of two possible physical solutions $t = (t_c, t_{\bar{c}})$, i.e. to remain conservative towards concessions ($t_{\bar{c}}$) or generally open to them (t_c), the US decided to accept the risk of an intra-alliance conflict and tacitly communicate the message that it would generally be open to making some concessions to North Korea. This outcome largely indicated what was to follow during the US-North Korea negotiations. If the South Korean utility vector is defined as $v = (v_{US}, v_{SK}, v_{Jap})$, and, on the other side, the utility vector derived from the American response $w = (w_{US}, w_{SK}, w_{Jap})$, then

$$\frac{w_{US} - v_{US}}{w_{US} - c_{US}} = r_{US}, \text{ with } r_{US} \geq p .$$

Herein, c_{US} is the conflict payoff that the US receives after rejecting South Korea's proposal, r_{US} is the highest probability of the conflict that the US would face rather than accept South Korea's offer. Finally, p stands for the probability that South Korea will not accept the US rejection as a new settlement. As the US was aware of how hard the negotiations with the North would be, it set its risk limit (r_{US}) quite high knowing that concessions would have to take place, while South Korea would hardly at that point accept the idea of substantial concessions. The US assumed the risk knowing South Korea would change its opinion in the mid-run. As already said, this calculation proved right.

Trying to mend fences after a period of disagreement, the Americans used a meeting in Washington to remind South Koreans that the June 11 statement was in English, not Korean, and that the US negotiators rejected a North Korean request that both sides should sign the statement. This was done in order to avoid any appearance of an official document. On June 25, President Kim would assess this meeting in Washington as positive, yielding benefits of "a balanced analysis". South Korean foreign minister would then publicly support the June 11 statement, adding that the use of sticks

and carrots alike was “necessary and right” (Wit et al., 2004, p. 68). Assuming the risk at the previous stage of intra-alliance bargaining proved correct.

Still, the South Koreans President kept swinging. He deviated from the previous stance once again on July 2, saying the North was abusing negotiations with the US to buy time necessary to finalize their nuclear weapons project (Wit et al., 2004, p. 68).

Coming soon after, from July 10 to 11, US President Clinton’s visit to South Korea had a soothing effect on slightly upset intra-alliance relations. He asserted that it would be pointless for the North to acquire nuclear weapons, “because if they ever used them, it would mean the end of their country” (Wit et al., 2004, p. 68).

Once the US-DPRK talks resumed in Geneva, the two sides agreed that North Korea would resume its negotiations with South Korea and the IAEA on the pace of prospective inspections by the Agency, and the US would, in principle, support North Korea’s conversation on light-water reactors that emerged as a possible solution to the crisis acceptable to all sides. However, the obstacle was the volume of financial commitments international community, and the US, South Korea and Japan more than others, would have to make in order to implement the project. Furthermore, this major investment required North Korea to make significant steps toward fulfilling its NPT obligations. This created the “who-goes-first” problem, and, more importantly, forged the problem of trust between the two sides. With respect to the intra-alliance relations, South Korea seemed generally satisfied with the Geneva talks’ outcome, as they seemed to have set the stage for inter-Korean dialogue, too (Wit et al., 2004, p. 75).

Hence, after South Korea returned to rather alliance-conditioned utility maximizing and calculating, alliance-conditioned utilities took over again. If the US agreement with North Korea is seen as an intra-alliance offer, then three physical outcomes are observed: $t = (t_a, t_r, t_a)$, i.e. accepting the agreement, rejecting it or trying to amend it, South Korea and Japan informally approved it, calculating their individual alliance-conditioned utilities as:

$$a_{SK}(x) = p_{US}\, p_{Jap}\, q_{NK}\ ;\ \text{and}\ \ a_{Jap}(x) = p_{US}\, p_{SK}\, q_{NK}\ \ \text{respectively.}$$

North Korea would, then, subsequently move to talks with the IAEA, whose director, Hans Blix, informed the North that without further Yongbyon inspections, the Agency would have to declare the “continuity of safeguards” broken. However, he was

trying to postpone the moment in order not to undermine the US negotiating efforts (Wit et al., 2004, p. 83).

At the same time, South Korea was trying to revive its own dialogue with its northern neighbor. South Korean Foreign Minister Han Sung Joo even suggested to the US Secretary of State Warren Christopher in July that the only way to push inter-Korean talks forward would be to make them a precondition of any better US-DPRK relations. Accordingly, Seoul devised a twofold plan linked to the progress of the IAEA-DPRK talks. In case Pyongyang agreed to an IAEA inspection in August, South Korea would propose negotiations in the framework of the Joint Nuclear Control Commission. If the North rejected this proposal, Seoul would follow a variation of the plan that would result in special envoys negotiations, appointing a more conservative envoy. Although the IAEA inspections started on August 4, the North rejected Seoul's initial proposal. Nevertheless, the South remained open to a special envoys exchange, suggesting a September 7 working-group level meeting to prepare an exchange of top leaders' envoys.

Consequently, as Seoul was getting more flexible, Pyongyang was becoming more rigid. It set two preconditions for talks: the South would have to show willingness to stop "nuclear war exercises" such as Team Spirit, and would have to refrain from working with other actors on sanctions against the North. Moreover, Pyongyang set September 20 as the deadline for Seoul to meet these demands. As this did not occur, another move had to wait until September 29, when Pyongyang proposed a session of talks with the South without prior conditions for October 5. Just before the session, Seoul decided to amend the original US-South Korean strategy which supposed trading Team Spirit for North Korea's commitment to stay in the NPT and accept international safeguards. Now it wanted to loosen the linkage, suggesting Team Spirit would be canceled if North Korean special envoy met President Kim in Seoul. After the White House first urged Seoul not to offer to cancel Team Spirit, and North-South talks made no essential progress, Washington decided to consider suspending the exercise if the North agreed to the special envoys exchange and IAEA's inspection activities. Although generally satisfied with this shift in the US attitude, the South preferred the exchange of special envoys as soon as possible, which was not likely, as it largely depended on progress in the US-North Korea negotiations. However, a secret memorandum written in

Seoul envisaged suspending Team Spirit when the North's envoy came to Seoul, even if that predated US-DPRK talks. Its clause never came to be activated as no progress on this issue was made. Furthermore, the North cancelled all the meetings with Seoul using as an excuse a statement given by the South Korean Defense Minister. Hence, the whole second track of the US diplomatic strategy collapsed. Accordingly, South Korea reinforced its demand to link any substantial US-North Korea talks to inter-Korean dialogue, a demand that would deeply complicate America's efforts to resolve the crisis before it erupted into an armed conflict. Aggravating the situation further on, it soon became clear that North Korea was not implementing the Geneva agreement from July 1993 (Wit et al., 2004, p. 88).

Another intra-alliance move that ensued was devising a "comprehensive approach", ambitiously offering a fourth round of negotiations to the North if it resolved a bulk of its nuclear issue at the next meeting with America's main negotiator, Robert Gallucci. As the impression within the Alliance was that step-by-step approaches largely failed, it was an alliance-level preference to turn to a rather "big package". Over time, it gained shape, entailing that North Korea would remain a party to the NPT, fully comply with its safeguards, and commit itself to implementation of the Denuclearization Declaration. In return, the US would conclude a peace agreement incorporating a legal guarantee on the nonuse of force against DPRK, assume responsibility for providing the North with the light-water reactor and normalize diplomatic ties with it. The key alliance element of this strategy was that South Korea and Japan were not to object normalizing US-North Korean relations as long as the nuclear dispute was solved first. On the other hand, international support for 'sticks' possibly coupled with 'carrots' of the big package was limited, even within the Alliance, as both Japanese and South Koreans preferred exhausting negotiations-driven diplomatic means first (Wit et al., 2004, pp. 96-101). At this point, individually defined preferences of South Korea and Japan on the nexus between the improving US-North Korean relations and intra-alliance abstinence in this matter played an important role, and were largely independent from intra-alliance considerations. South Korea's utility calculation took the following course:

$u(x)_{SK} = f(p, q, m)$, or in an equation: $u(x)_{SK} = pqm$.

Herein, p is the probability that the US will keep linking inter-Korean dialogue to the US-North Korean talks, q the probability that the US will not assume any new responsibilities on behalf of South Korea as a part and precondition of the improving US-North Korean relations, and m stands for the probability that North Korea will fulfill its obligations. For Japan, on the other hand, this utility function is simpler, as North-South talks play no role in its direct calculations and can, therefore, be abstracted from. Given the above conditions stated in the utility functions described, and given the quasi-simultaneity of events at this point, a payoff matrix can be designed. Strategy choice for the US is to normalize relations with North Korea as it starts fulfilling its obligations, or not, therefore making a strategy set $S_{US} = (N, \overline{N})$. For South Korea and Japan, on the other hand, choices are to support the progress in US-North Korean relations, or not, therefore $S_{SK,Jap} = (A, \overline{A})$, as 'approve' and 'not approve'. Given the description of probabilities p, q and m, and the US commitment to North Korea after Geneva, then the following matrix is designed:

SK, Japan

US /	A	$\overline{A}$
N	**4,4**	**3,1**
$\overline{N}$	**1,1**	**0,0**

Table 1. US-North Korea negotiations approval matrix

Cardinal values in the matrix are not derived from the utility functions depicted above, as it is at this point of less relevance, but are meant to clearly indicate incentives for specific strategy choices. As depicted, A strictly dominated $\overline{A}$ on the part of South Korea and Japan, and $\overline{N}$ was strictly dominated by N on the part of the US. Therefore, the strategy profile $s_{US,SK,Jap} = (N, A)$ was a unique Nash equilibrium, which indicates the game involved a range of strictly converging intra-alliance interests, determined by the US commitments to North Korea and relatively high South Korean and Japanese expectations on the values of probability factors p, q and m defined earlier.

However, pressed by domestic pubic and hardliners within his own government and bureaucracy, South Korean President Kim took a tougher position right before the summit with the US President Clinton in Washington. The summit was scheduled for November 23 and the Americans expected no major surprises. Preparing the décor for the summit, they told the South Koreans that their patience with the North was diminishing but also that the US would stay open to constructive negotiations while implementing the comprehensive approach commonly agreed upon within the Alliance (Wit et al., 2004, p. 109).

What stunned them was President Kim's idea that the US and South Korea should cancel Team Spirit only after the North had demonstrated to the South it did not have any nuclear weapons. Counting on diplomatic games and the 'dignity argument' by North Korea, the Americans saw this demand meant setting the bar too high, probably leading to a complete deadlock (Wit et al., 2004, p. 111).

The Americans tried to convince the South Korean president not to abandon the comprehensive approach, which finally made him conclude that the problem with this approach might have been just one of terminology that possibly deluded the media. Finally, the approach was renamed to "thorough and broad". Additionally, the allies agreed that Team Spirit could be canceled only after a special envoy of the North had visited Seoul and had held "serious talks". In conclusion of the summit, one South Korean newspaper made a remark that the President Kim Young Sam said 'no' to an American president, something that preceding South Korean leaders had been unable to do (Wit et al., 2004, p. 113).

Therefore, a shift in South Korea's preference took place due to significantly irrational factors, deviating from the previously agreed course of actions. After the South Korean proposal to cancel Team Spirit only after the North had proved it had no nuclear weapons, the possible physical outcomes were for the US to accept this demand (t_a),or try to reverse it by proposing the " thorough and broad" mantra (t_r). Accepting South Korea's demand would result in the utility vector $v = (v_{US}, v_{SK}, v_{Jap})$ with divergently perceived and disparate benefits whereas the alteration of it would result in an alliance-conditioned utility a, approximate for all the three allies. Again risking an intra-alliance conflict due to a *de facto* rejection of the South Korean proposal as part of the altered US

offer, the US opted for a modification and proposed the "thorough and broad" approach. Again, the US accepted the risk of intra-alliance conflict, knowing the South Korean demand was impossible to propose as an alliance-level preference. Therefore, r_{US}, defined as $\frac{a - v_{US}}{a - c_{US}}$, was again higher that the probability that South Korea would reject the American modification of their demands (p). And again their calculations proved right, as it was demonstrated above through the remark of the South Korean president at the end of the summit.

On December 29, a preliminary agreement between the US and North Korea was reached. The two sides adopted the principle of "simultaneous and reciprocal actions". They would take four steps on the same yet unspecified day, in advance dubbed "Super Tuesday" by the Americans. North Korea would start receiving IAEA inspectors at seven concrete sites, would resume dialogue with South Korea to arrange the exchange of special envoys, whereas Seoul would announce the cancellation of Team Spirit and the US and North Korea would specify the date for a meeting between the North's and the American key negotiator, Robert Gallucci. The agreement was endorsed by South Korea, yet their concern about the specific timing of the special envoys exchange remained (Cronin, Bowman, Collier, Niksch & Shinn, 1994; Wit et al., 2004, p. 116).

Nonetheless, the beginning of 1994 saw an exacerbation of the situation. While the IAEA-DPRK talks halted, the press got a classified Pentagon's document on the US plans to deploy Patriot antiballistic missiles in South Korea. This new development caused serious debates in South Korea, dividing the country internally and raising doubts the US might be jeopardizing South Korea's national security by deploying the Patriots, or even trying to get the South to purchase them on a commercial basis, not under the regular and less costly foreign military aid program. Other reports followed soon: on the gist of the American war plan, emphasizing its goal to seize Pyongyang, and on Pentagon's plans to send additional 1.000 soldiers to Korea if the Vienna talks failed. Exacerbating the situation even more, the US Senate passed a tough-worded nonbinding resolution, one advocating sanctions against the North and the other calling upon the President to consider redeploying American nuclear arsenal in South Korea. Reacting to these developments, South Koreans started fearing the US might take them unwillingly to

war. They also brought the South Korean government into trouble as it all looked like a deliberate attempt to challenge the North. Overtly, however, the government said that was a proof of the US commitment to Seoul and South Korean security, adding that it had much appreciated the missiles. On the other hand, South Koreans covertly maintained and finally communicated to the press that no American fait accompli will be accepted and that the deployment of the Patriots was still in the phase of discussion. They also knew that the US President Clinton approved the deployment to begin at the end of February 1994 (Wit et al., 2004, p. 125).

The press releases perilously severed the US-North Korean relations, as the North thought US measures were a dress rehearsal for a second Korean War. Consequently, North Korean rhetoric against the US and Patriot missiles promptly escalated. After a brief progress in negotiations and North Korea's acceptance of the IAEA's terms on inspections, another setback followed. The North suddenly shifted gears, saying the Agency's inspection would not be possible before a major progress in its talks with Washington is achieved. Adding to the crisis, South Koreans remained unyielding as well, averring the special envoys exchange had to take place before any agreement between the US and North Korea. However, in yet another twist, new agreement emerged in February 1994, repeating the provisions of the December 29 accord and specifying a new round of talks would start on March 21. Supporting the emerging solution, South Korean President Kim announced he was ready to meet Kim Il Sung. North Koreans, conversely, ridiculed the proposal, calling Kim a "puppet". What the Americans called 'Super Tuesday' was announced on March 3, 1994 (Wit et al., 2004, pp. 126-139).

Exactly one year after North Korea's announcement of its withdrawal from the NPT, Robert Gallucci, the leading US negotiator, visited Seoul while working on the implementation of Super Tuesday. During one of his meetings with ROK's officials, he compared the US to a tugboat that had to keep pushing two barges simultaneously, the IAEA and South Korea, towards North Korea in order to note any progress. Negotiations with North Korea, on the other hand, started on March 3 in Panmunjom and went back and forth. What was important for intra-alliance relations was an agreement that there could be no solution to the nuclear crisis without North-South dialogue. If those talks

failed, the Patriots should be deployed and Team Spirit reinvigorated (Wit et al., 2004, pp. 143-146).

Trying to move ahead with their bilateral talks with DPRK and honor their promise to the South that a new round of negotiations with the North would not take place before the special envoys between the two Koreas had been exchanged, the Americans came up with a logical solution: to perform these operations simultaneously. South Koreans remained generally flexible. The intra-alliance game was at that point turning into one of alternating offers solved at the first stage of bargaining, right after the American suggestion (Wit et al., 2004, p. 147).

In a quickly changing situation, the IAEA inspectors left North Korea with the job unfinished, as their work was thwarted by the hosts. A series of tough stances exchanged between North Korea and the US followed, undermining prospects for the success of Super Tuesday. Soon after, the North gave an ultimatum to Gallucci, saying the envoy exchange would have to take place only after a new round of US-DPRK talks, and if Washington tried to evade its obligations to cancel Team Spirit and hold talks with DPRK, it would withdraw from the North-South dialogue and suspend its cooperation with the IAEA. They gave Gallucci three days to respond, but the response was equally tough, stating Pyongyang still had to complete the inspection and the exchange of envoys before March 21, which was the date for a new round of the US-DPRK talks. The climax of that diplomatic downfall was the North-South talks March 19 session, when the North warned its southern neighbor that if war broke out, there would be “a sea of fire”, as Seoul is very close to the border. After press leaks, the world mostly blamed the North for the diplomatic failure, strengthening the American position within the UN and expanding the support for sanctions against the North (Wit et al., 2004, pp. 147-149).

Consolidating intra-alliance views, the US and South Korea agreed that if North Korea failed to implement Super Tuesday conditions, Team Spirit would be rescheduled and Patriot deployments would proceed. Moreover, South Korea preferred an urgent shipment of the Patriots, which took place soon afterwards and first missiles were in South Korea on April 18. Super Tuesday was de facto dead and the Americans learnt they could not let their junior ally (South Korea) determine the pace and destiny of negotiations anymore (Wit et al., 2004, p. 152).

This series of adoptions of South Korea's intra-alliance proposals by the US and of a pronounced deference to South Korea's preferences can be qualified as the South Korean invitation to the US to ally more closely to it. Although such an offer had never been made, it fruitfully sums up a chain of events that followed the identical path of the American intra-alliance concessions to South Korea. South Korea's proposal produced a set of physical outcomes $t = (t_a, t_r)$, with the US accepting the South Korean proposal or rejecting it. Given the rising rigidity and frustration with the South Korean leadership, the Americans thought ignoring their efforts for a tighter alliance would be interpreted as an act of sidelining and possibly even cheating on the allies while negotiating with North Korea. Having defined the alliance-conditioned utility at this point as:

$$a(x) = f(V_{rel}, p),$$

where:

$$V_{rel} = \begin{cases} V_{abs}(1+m), & m \in [-1,0) \\ V_{abs}, & m = 0 \\ V_{abs} + (1 - V_{abs}) * m, & m \in (0,1] \end{cases}.$$

Herein, V_{rel} is the relative value of the alliance for each player, V_{abs} is the absolute (overall) value of the alliance for each player, with $V_{abs} \in [0,1]$; and m stands for the *issue-area index*, essentially a ponder indicating the cardinal value of an alliance priority once the salience of alliance preservation (previously defined as 'absolute value of alliance') faces particularistic interests of specific allies in a concrete situation. The overall interval of m is $[-1,1]$, where $m \in [-1,0)$ in cases of a declining priority of the alliance, i.e. its negative valuation in a specific issue-area, and $m \in (0,1]$ in cases of a rising priority of the alliance, i.e. its positive valuation in a specific issue-area. Finally p stands for the probability of North Korea favorably concluding the negotiations, thus practically indicating each ally's confidence about the pace of negotiations. As the alliance-conditioned utility was at this point formulated only by the US, and as it assigned an increasing priority to the alliance, along with the faith that its negotiations with North Korea were starkly progressing, the American utility function was:

$$a(x_{US}) = V_{rel}\, p, \text{ with}$$

$$V_{rel} = V_{abs} + (1 - V_{abs})m .$$

Evidently, the US believed that the conflict payoff in the case of rejection of the South Korean demand was rather low, i.e. that the cost was too high. Therefore, it kept appeasing the South, which proved to be a flawed strategy once Super Tuesday failed. Observing a fictive decision-making process by the US through the above formulas, it can be concluded that the Americans overestimated two variables: the alliance priority, or more precisely the issue-area index (m), but even more the pace of negotiations with the North (p), having too much trust it would honor the preliminary agreements. The reality was a bit different, saying $a_{US} \leq c_{US}$, which meant that the benefits of alliance were at that point lower than the costs.

The gradual escalation strategy was revived, as the Transpacific Allies agreed to give diplomacy yet another chance, especially as the IAEA did not find the continuity of safeguards broken. The UN circle of friends, partners and traditionally tough states was in the game again, as the US was trying to get the Security Council members on board for a resolution introducing sanctions to the North. Faced with the cancellation of a new round of talks with the US, the Patriot deployment, rescheduling of Team Spirit and potential UN sanctions, North Korean foreign minister threatened again, stating his country might withdraw from the NPT (Wit et al., 2004, pp. 156-159).

On March 31, 1994, UN Security Council issued a unanimous presidential statement calling upon Pyongyang to allow to the IAEA inspectors to complete their inspections in accordance with the February 15 agreement. In response, North Korea signaled more talks were needed (Wit et al., 2004, pp. 160-161).

Another visit of Gallucci to Seoul was meant to foster intra-alliance support for de-linkage of further US-DPRK talks from the exchange of special envoys. Fearing this mutual conditioning could lead to a disaster of negotiations altogether, the South decided to compromise by agreeing an inter-Korean special envoys meeting would take place on the first day of the new round of the US-DPRK negotiations (Wit et al., 2004, p. 169).

After the North warned it would defuel its 5-megawatt reactor due to allegedly technical reasons, the Alliance quickly realized: this would bring North Korea much closer to producing five or so nuclear weapons. This reality prompted the US military contingency planning in and with Japan. In the same spirit, US State Secretary Perry told

the South Koreans that the Alliance and the US more specifically would not initiate or provoke the war, but should not invite it either through lack of preparation. This was meant to sober up the South Koreans, as they kept on doubting the American contingency planning as a self-fulfilling prophecy that, at the end, might lead to an actual armed conflict (Wit et al., 2004, pp. 171-179).

Facing the peril of defueling, the Security Council issued a presidential statement on May 30, 1994, inviting DPRK to proceed with the defueling only under the IAEA's supervision in order to avoid misusing the nuclear material. The Security Council urged North Korea to leave room for fuel measurements that would eventually give data on the history of the 5-megawatt activity and reveal whether cases of misuse had already taken place in the past (Wit et al., 2004, pp. 185-189).

At the same time, Northeast Asia was relapsing into a full-fledged crisis: only a day after the SC statement, North Korea tested an antiship missile in the Sea of Japan, which was interpreted by experts as a an implicit threat to Japan to refrain from supporting sanctions within the UN. The American response was harsh, restating the military commitment to the security of its allies and reconfirming the intra-alliance solidarity (Wit et al., 2004, p. 189).

However, by unloading the 5-megawatt reactor, North Korea destroyed historical information on its previous nuclear program. Washington responded by canceling talks with Pyongyang and pressed the UN to impose sanctions on the North. South Korea and Japan followed. South Korea even imposed its own unilateral economic sanctions on its northern neighbor. International support for sanctions was growing and even China seemed to begin showing some flexibility faced with North Korea's possible armament efforts. Diplomatic campaigns of the US and North Korea even clashed in the Non-Aligned Movement. Yet, the Americans won the majority support for its efforts there, too (Wit et al., 2004, pp 196-200).

On the other hand, South Korea was initially skeptical towards the idea of the former President Carter that he might be the turning point and that he would bring a diplomatic solution about. Seoul thought Carter's personal diplomacy would marginalize its role again. Therefore, it was not surprising when the South Korean President Kim told his American counterpart Clinton that Carter's forthcoming visit to North Korea was a

mistake, especially as international support to sanctions was growing. Still, Seoul avoided hinting publicly that it opposed the visit (Wit et al., 2004, pp. 202-204). This decision was a result of a backward induction-like reasoning of South Korea. It could construct a game tree without ever playing the game with the US just by following the large stock of information about other allies' preferences due to intensive intra-alliance information exchange and consultations.

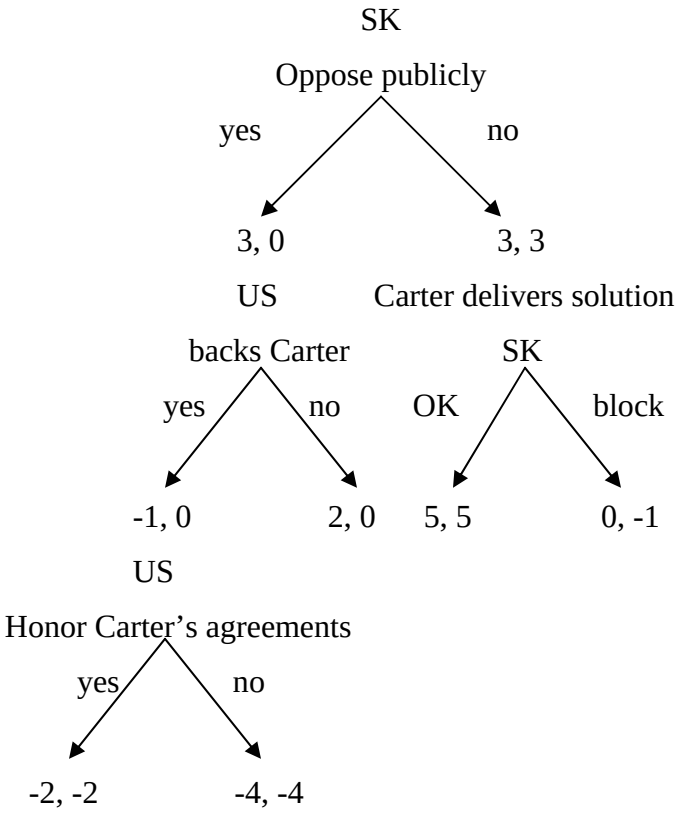

Figure1. South Korea's projected game tree

As depicted above, South Korea could publicly oppose Carter's mission to the North or not oppose it publicly. Open support was not an option. In case it opposed it, the US government would then move next. They could either back former president Carter at that point or stay idle. In case the US backed him, South Korea would slightly lose on the international stage and within the alliance, *de facto* losing US support in that matter. In

case the US did not support Carter, South Korea could win, but it knew that was not a particularly probable outcome. On the other hand, by practically forcing the US to support Carter, South Korea would unintentionally cause Carter's mission to be regarded as *de facto* official by North Korea. Then, any serious talks with the North would be probably impossible, and if Carter was to reach any agreement with the North, it would be probably unfavorable to both the US and South Korea, given Carter's highly critical statements both before he left for North Korea and after he met President Kim Il Sung. Yet the US would have to honor it. In case it decided not to honor it, the damage would be even greater in short, mid and long run. After having assigned probability values to each vertex of the tree, the South Koreans could see that the end point would probably be the US being obliged to honor Carter's agreement with all the negative payoff it would bring along to both the US and South Korea.

On the other hand, South Korea could decide not to oppose Carter's mission publicly, thus allowing for the option of Carter either reaching a solution favorable to them or not. This time, prospects for a generally favorable solution were much higher, as he would not be acting as a US official. Then, if South Korea liked the solution, it could always approve it and possible win significantly much. If it did not like it, it could block it within the alliance with a very high probability. Cutting off the ending branches one by one, South Koreans made the right decision, which was soon to be proven.

US commanders were simultaneously discussing a military buildup in South Korea to secure their forces and prepare for a potential war in case the North does interpret prospective UN sanctions as a *casus belli*. The South Koreans, conversely, were reluctant to allow a rapid buildup of US forces on the Peninsula, believing it would only precipitate a North Korean attack. Moreover, American military planners and the US ambassador to South Korea agreed that the US civilians and other foreigners should be evacuated before the imposition of UN sanctions on North Korea. It was not clear whether the US allies, South Korea and Japan, would support the Osirak option, i.e. a preemptive strike on Yongbyon facilities similar to what Israel did in Iraq in 1981. However, the Allies did support the US draft sanctions resolution on North Korea (Wit et al., 2004, pp. 210-212).

Pressured, nevertheless persistent, the North Koreans were sending mixed signals, first stating they would be ready to discuss all matters of American interest, and then saying sanctions would mean war, and there could never be mercy in war (Wit et al., 2004, p. 213).

Around the time of Carter's arrival in Seoul on June 13, the US and South Korea were closely working on preventing a spontaneous and disorderly evacuation that might have taken place, and on arranging the planned one, ensuring, for instance, that trains were available to transport Americans south to embarkation points. Nonetheless, the South was suspicious that the Americans might have been working behind their back and were seriously exploring the military option. There was a problem of trust within the alliance. The South Koreans even said they "were not ready to be sacrificed on the altar of nonproliferation". Years later, President Kim would claim the US was seriously about to announce the evacuation of its citizens and ignite a conflict with the North, without even consulting its ally. He also claimed that he held Washington back through his timely interventions with the US ambassador to South Korea and President Clinton. On the other hand, there is no official record of any communication at this time during which President Kim expressed to President Clinton his concerns about the evacuation or war (Wit et al., 2004, pp. 217-220).

Before Carter left Seoul, President Kim had given him a straightforward proposal that the leaders of North and South meet without any prior conditions (Wit et al., 2004, p. 221).

As Carter's diplomatic activity in Pyongyang and his controversial CNN statement had quite some resonance with the North, the South Koreans feared they might be sidelined again in that diplomatic struggle along with being concerned Carter's personal diplomacy might undermine the increasing international pressure on the North. Carter's appeasing statements towards North Korea and his allegations that the US policy on the matter was deeply wrong and ineffective brought intra-alliance relations to the fracture point. South Koreans did not really understand what was going on, but firmly believed the Americans sidetracked from the common stance of the Alliance. Still, President Clinton managed to calm his South Korean counterpart, and it became all easier once Kim Il Sung accepted an inter-Korean summit. This instantly gave a strong

credibility to Carter's mission on the part of South Koreans (Wit et al., 2004, pp. 231-235).

The result of former president Carter's visit to North Korea from June 15 to 18 was that North Korea committed itself to let the IAEA inspectors and equipment remain at Yongbyon, and to arrange an inter-Korean summit between South's Kim Young Sam and North's Kim Il Sung. Soon afterwards, the US and North Korea agreed through the New York channel to resume US-DPRK talks on July 8. On June 28, the North-South Korean summit was scheduled for July 25-27 (Wit et al., 2004, p. 240).

The launch of negotiations was marked by Kim Il Sung's sudden death and a lukewarm South Korean reaction coupled with a doubt he might have even been killed by insiders. Moreover, a theory of North Korea's early collapse reemerged now that DPRK was left without the 'Great Leader'. About the same time, Seoul released documents proving beyond doubt Kim Il Sung's responsibility for the outbreak of the Korean War. This chain of events infuriated the North, calling for urgent US damage-control measures in order to save the negotiations that had barely started. Accordingly, the Americans accepted a North Korean invitation for Gallucci to pay a condolence call at their Geneva mission and sign the condolence book contrary to the South Korean advice. This gesture would be deeply respected by the North Koreans later on. What really mattered for the US was no change in North Korea's negotiating position (Henriksen & Mo, 1997, pp. 2-12; Wit et al., 2004, p. 259).

However, Kim Il Sung's death did not fail to cause yet another intra-alliance rift. Frustrated as the Americans did not accept their suggestions on condolences, Seoul believed this gesture would only encourage hardliners in Pyongyang. Furthermore, the Americans were not in a wait-and-see mood. They did not think, quite differently from the South Koreans, that the time was on their side. Washington thought resolving the nuclear problem could not be delayed, even with the Great Leader's death playing role. Seoul believed it could no longer influence the American decisions whatsoever, even publicly complaining their allies were too hasty. Seoul's complaints about its diplomatic exclusion were renewed (Wit et al., 2004, p. 265).

Having factored South Korea's demands into their utility calculation again, the Americans must have reached a different result. With the experience of Super Tuesday

failure, and given the delicacy of the moment after the Great Leader's death, this time, their utility calculation was in principle still alliance-conditioned, however more realistic.

It was:

$$a(x_{US}) = V_{rel}\, p \text{, with } V_{rel} = V_{abs}(1+m);$$

and $m \in [-1,0)$, as the alliance priority was declining.

In this case, too, p is the level of confidence about the outcome of negotiations with North Korea. However, not overestimating the issue-area index (m) this time and still assigning a relatively high value to *p*, the US concluded that $a(x_{US}) \leq c_{US}$, i.e. the conflict payoff was then higher than the alliance-conditioned utility. So it went for a conflict.

Nevertheless, the US multilateral diplomatic approach had to move on. Gallucci was set in July 1994 to build an international support for the light-water nuclear reactor project. Convincing the South Korean ally took assuring it that its light-water model would be the base of the North Korean one and that South Korean companies would play a major role. Along with opportunities, possible pitfalls came, too. South Korea would have to finance the most of the project and the question of political viability remained, as it was clear from onset that insulating the project against the ups and downs of inter-Korean relations was a major challenge. Nonetheless, President Kim thought the project could be a chance for an inter-Korean economic rapprochement, a possible interlude to reunification. Finally, he supported the idea as a long-term opportunity with multiple benefits (Wit et al., 2004, p. 267).

The Japanese third of the Alliance was less enthusiastic. They were initially ready to financially support some sort of feasibility study or a survey of the North's energy needs. However, a multibillion dollars commitment had to wait for improved political relations with the North. The Americans, on the other hand, concluded there was still some bargaining room with Tokyo, especially if a comprehensive solution to the nuclear issue was reached. The US did essentially manage to secure a general yet vague international support for what Robert Gallucci labeled KEDO, the Korean Peninsula Energy Development Organization. It will be an international consortium in charge of constructing the North Korean light-water reactor. After days of negotiations, North Korea said OK to the international consortium idea, agreeing with the US its company

will be a prime contractor. The US idea, however, was that the real construction and implementation of the project would be conducted by a South Korean company, a formal sub-contractor (Wit et al., 2004, pp. 268-275).

In a one-week period, in August 1994, the North Korea and the US signed the Agreed Statement showing substantial progress toward a comprehensive settlement. Therein, the US committed itself to provide light-water reactors to North Korea and make arrangements for interim energy supplies, while North Korea promised to remain in the NPT, allowing implementation of its safeguards agreement and implementing the Joint Denuclearization Declaration signed with the South. The Statement also envisaged a gradual establishment of diplomatic ties between the US and DPRK. And although the Statement did not specify that, the North Koreans seemed open to South Korea's participation in the reactor project (Wit et al., 2004, p. 277).

South Korea's reaction was lukewarm. It feared Pyongyang was taking advantage of the American misperceptions and excessive enthusiasm. They felt that the Americans had committed the entire Alliance without getting any tangible North Korean commitments in return. President Kim even repeated to the press what he had previously said to President Clinton:" Excessive progress is being made in US-North Korean talks despite the unstable North Korean situation". Nevertheless, the South Koreans yielded to the American persuasions, convinced by their arguments on the grand opportunity that was appearing, but more than anything, faced with an American beefed up diplomacy that showed no intention of waiting for the South. The Americans had been repeatedly assuring the South Koreans that North-South talks remained a policy priority for the US. However, a sort of intra-alliance fait accompli and the sped up US-North Korea talks confronted the South with a multilateral diplomatic inertia it could not resist. Finally, the South Koreans did believe the light-water reactor might be a channel through which the two Koreas could reconnect. As for the trust, they somehow accepted that if the Americans decided to trust the North, they would have to trust them as well. After all, South Korea's homeland security did depend on the US. On October 13, the South Korean leader was ready to give in and support the Geneva agreement. As the South's foreign mister put it, the president "was ready to be a statesman". The deal as sealed by President Clinton's call to President Kim, when the South Korean leader emphasized a

US public restatement of its security commitment to South Korea was needed. On October 15, South Korea publicly signaled its acceptance of the probable US-North Korean agreement. The support was evidently lukewarm. What encouraged the South was Gallucci's October 15 insistence not to drop the issue of North-South dialogue during his negotiations with the North Koreans. He made it an integral part of the immanent deal (Wit et al., 2004, pp. 280-289 & pp. 319-323).

About the same time, Gallucci managed to get the Japanese on board as well. They found the idea of a multinational consortium safe and beneficial enough; especially after Gallucci convinced them that the North Koreans were serious about the resolution of the crisis through the light-water reactor gateway. Negative security assurances that were about to be given to the North seriously worried Japan, as it thought assuring North Korea it would not be attacked by nuclear weapons undermined its own security. However, the Americans again assured the Northeast Asian ally that its security commitment was unshakable, and would be especially relevant if North Korea failed to comply with its NPT obligations in future. Additionally, the Japanese believed this project might bring them closer to the South Koreans. Tokyo was fully aware of intra-alliance tensions, and was not any less skeptical towards the Americans than Seoul. Japanese diplomats thought a rapprochement with Seoul might help them build a common intra-alliance front against the US and decrease their political inferiority. Finally Tokyo generally accepted to commit financially, too (Wit et al., 2004, p. 290).

It was evident that, as depicted in the previous passages, the South Koreans and the Japanese turned to the alliance-conditioned utility calculating. Their utility function at that moment could be defined as following:

$$a(x_{SK,Jap}) = V_{rel} q_{NK} i * n ,$$

with

$$V_{rel} = V_{abs} + (1 - V_{abs}) m ,$$

and $m \in (0,1]$.

This utility function from accepting the outcomes of the US-North Korean negotiations was determined by the high relative value that the South Koreans and the Japanese

assigned to the Transpacific Alliance at that point, knowing that at a moment when big agreements are being made, they had to reaffirm their rights and the US security commitments. And although they did not really believe this effort might resolutely change the North Korean attitude (low q_{NK} probability), they did not want to remain isolated on this issue, which was very probable (high ponder *i*), and were indeed aware that the project, if successful, might bring them an enormous political and economic gain, following possibly closer relations with the North. Although the factor of future prospects existed, not much confidence was related to it (relatively low *n*).

On the other hand, rejecting the prospect of $a_{SK,Jap}$ induced by the American offer of the agreement reached with the North, South Korea would have to accept the disagreement value, i.e. the conflict payoff $c_{SK,Jap} = 0$. The agreement point of the game was the unique alliance-conditioned utility vector $\bar{a} = (\bar{a}_{US}, \bar{a}_{SK}, \bar{a}_{Jap})$, which maximized the product

$$\pi = (a_{US} - c_{US})(a_{SK} - c_{SK})(a_{Jap} - c_{Jap}).$$

This agreement point was the alliance-level decision to back the agreement, delivering Nash product to the bargaining, as with all the three allies

$$a_{US,SK,Jap} \geq c_{US,SK,Jap}.$$

The signing ceremony of the Agreed Framework between the United States of America and the People's Democratic Republic of Korea took place in Geneva, on October 21, 1994. It stated the US and North Korea would work together to replace North Korea's graphite-moderated reactors and related facilities with light-water reactors. Furthermore, it promised normalizing US-North Korean political and economic relations as progress is made on the nuclear issue, and committed both sides to work together for peace and security of the Peninsula, which entailed inter-Korean dialogue, too. Finally, the North assume the obligation to remain a party "in good standing" to the NPT, allowing implementation of its safeguards. It specified that, for North Korea traditionally problematic, special and ad hoc inspections by the IAEA would resume once the supply contract for the provision of the light-water reactor (LWR) project is concluded. North Korea accepted to come into full compliance with its safeguards agreement with the

IAEA after a "significant portion of the LWR project is completed, but before delivery of key nuclear components" (Gilinsky, 1997, p. 10; Wit et al., 2004, p. 329).

The response in South Korea and Japan was on the overall a positive one, yet cautious. Both accepted to proportionately fund the LWR project, aware they would have to make a great deal of effort to persuade their respective publics in spite of the critics. This was due to the Northeast Asian history of zero-sum games. For the South Koreans and Japanese, it was hard to believe that all of the sudden a solution that was good for them might be good for Pyongyang, too. Following a mistake in President Clinton's weekly radio address at the beginning of November 1994, the South Koreans got strongly agitated as they thought the US had imposed another burden upon them, paying for oil shipments to North Korea, without their prior consent. After a clarification given by the US State Secretary and his explanation that a well-crafted point was wrongly communicated to the public, South Korean complaints were toned down. Soon afterwards, controversies on the decision-making structure within KEDO and the role of its executive director[119] were resolved, too, allowing for the project's kick-off (Wit et al., 2004, pp. 344-350).

5.2.2 Second Korean crisis: October 2002- February 13, 2007

Eight years after signing the Agreed Framework, the US Intelligence discovered a secret North Korean program to produce highly enriched uranium (HEU) for nuclear weapons apart from the plutonium production frozen under the Framework. What caused this stunning disclosure was a series of events that made the North feel both militarily and economically insecure. Moreover, the North Koreans never stopped believing their nuclear program was their one and crucial bargaining chip, the element whose absence would render North Korea just another poor country of nobody's interest. Its possession, on the other hand, made it one of the top US security priorities and a constant concern of South Korea, Japan and China. After years of floods, droughts and famine in North Korea, the spy submarine incident of 1996 in South Korean waters, the June 2002 North-

[119] Which was planned to be an American, thus completing the picture of the US leadership in the project, needed to 'buy' the approval of North Korea

South naval vessels gun battle, and the slowdown in the LWR construction coupled with KEDO's indebting and financial insolvency, Pyongyang started believing its core means of survival was not a dependence on foreign aid or agreements, but a nuclear self-reliability that, quite contrary, meant violating those agreements. This, however, was not an entirely true picture, as the North did very much count on foreign, especially Chinese and South Korean aid and foreign direct investments to beef up its shattered economy. Keeping tensions over its role in Northeast Asia alive, North Korea test-launched its Teapodong-1 long-range missile in August 1998. Bypassing the fact that the missile flew over the Japanese island of Honshu, North Korea focused on denying this occurrence was meant to be perceived as its attempt to intimidate regional players and gain yet another bargaining leverage. Yet, it simply claimed it was an unsuccessful satellite test (Wit et al., 2004, pp. 372-374).

Anyhow, its HEU program was definitely more dangerous, as the technology required to design a uranium bomb was much easier to master than that required to build a plutonium bomb. And although the Agreed Framework did not mention bans on HEU programs, it did refer to the North-South Denuclearization Declaration prohibiting this and all related activities. So, the North Koreans did cheat. Moreover, in the eyes of the US, they had already become a piece of the 'Axis of Evil' in the aftermath of 9/11, and the stage was set for yet another round of enmity. The US claimed that Pyongyang admitted it had been developing an indigenous HEU program, whereas North Korea denied such allegations. In November, the US persuaded its KEDO partners, including Japan and South Korea, that heavy oil deliveries to Pyongyang should be suspended. In December 2002, North Korea responded to these new developments by expelling IAEA inspectors monitoring the nuclear freeze, reloading and restarting its 5-megawatt reactor in Yongbyon. Finally, on January 9, 2003, North Korea withdraws once again from the NPT (O'Hanlon & Mochizuki, 2003, p. 32).

By this point, diverging views within the Transpacific Alliance on the reemerging nuclear crisis had already appeared. After Japanese Prime Minister Junichiro Koizumi's historical visit to Pyongyang, somehow similar to the North-South summit of 2000, this crucial US ally would significantly change the course of its policy towards North Korea. Its policy of opening towards the 'Hermit Kingdom' was frustrated as North Korea's Kim

Jong Il only ostensibly (at least as the Japanese saw it) regretted the abductions of Japanese citizens in the 1970s and 1980s, astonishing the world, yet only confirming what the Japanese had been suspecting ever since. Since then, Japan has been pursuing the so-called 'Libyan model' for the solution of the North Korean nuclear issue, claiming Pyongyang should first freeze and dismantle its nuclear program, and only then should the international community take it back into the system, developing political and economic relations with it. In the meantime, North Korea should be exposed to a limited containment policy, obstructing its access to international political, financial or economic support. Possible military buildup within the limits of Japan's pacific constitution was widely seen as needed, too, for self-defense purposes. Japanese adherence to the Libyan model practically brought them closer to the US, whose administration insisted on following the identical path. This, however, did not mean adopting the same American posture of preemption (Park, 2005, p. 85).

On the other hand, Bush tuned a cold shoulder to South Korea's President Kim Dae Jung in 2001, known for his "Sunshine Policy" of active engagement towards the North, similar to what Japan shortly tried to do before its Prime Minister's visit to Pyongyang in 2002. Once it became clear that North Korea breached its nonproliferation obligations, the US stance became even tougher, as its diplomats started pursuing a new string of sanctions towards Pyongyang and reaffirming the military option in dealing with the North. South Korean response to this American diplomatic offensive was cautious. Its newly elected president, Roh Moo Hyun, had previously dismissed the military option categorically and kept refusing to support it even after President Bush publicly hinted in March 2003 it might be on the table. Evidently, the US proposal of preemptive strike philosophy over North Korea, and the South Korean silent refusal to second it, deteriorated their intra-alliance relations. Divergence was clear as President Roh Moo Hyun kept roughly following former President Kim's "Sunshine Policy", now dubbed "Peace and Prosperity Policy" (O'Hanlon & Mochizuki, 2003, p. 41; Park, 2005, p. 80).

Soon afterwards, in an effort to repair relations between the allies, President Roh convinced South Korea's national Assembly to approve deployment of non-combat troops to support the US-led intervention in Iraq at the end of March 2003. However, the effect remains limited due to a rather longstanding, yet vague mistrust the South Koreans

have been feeling for the US in the past five decades (O'Hanlon & Mochizuki, 2003, p. 41). Dilemmas remain on how far the US would go to contain the North and keep it from attaining nuclear weapons, and if South Korea would be sacrificed by the US on the altar of nonproliferation. Also, a growing anti-imperialist sentiment towards the US should not be ignored either.

Another reason why the fence was not really mended is quite simple: South Korea did not change its opinion on the policy of coercion seen as a sword of Damocles hanging over North Koreans' heads. This stance was reinforced during President Roh's visit to Washington in May 2003. He kept advocating a rather engaging posture towards DPRK (O'Hanlon & Mochizuki, 2003, p. 84).

Trying to bridge the ever growing gap between the US and North Korea, the Chinese organized a three-party meeting in Beijing hosting the US and North Korean diplomats. Much to the disappointment of South Korea, the US delegate barely repeated the Administration's declared position that North Korea must dismantle all forms of its nuclear program before the US would be willing to discuss ways to improve the US-North Korean relations. What happens from that point onwards is know as the Six-Party Talks involving transpacific allies (the US, South Korean and Japan), North Korea, China and Russia. Also first suggested by China, the talks have become a key framework for US-North Korean negotiations, bypassing the troublesome bilateral format that the Americans deemed unacceptable before North Korea returned to the NPT (Niksch, 2004, p. 6).

The specificity of the Six-Party Talks with respect to the US-South Korean-Japanese intra-alliance relations is a bit paradoxical. Although now actively participating in negotiations with North Korea, contrary to the 1993 and 1994 negotiations, South Korea and Japan remained on the sidelines. Firmly believing in a distinct approach to the North, South Korea kept advocating policy softening and engaging North Korea diplomatically, pushing for the so-called 'Ukrainian model' towards its northern neighbor (Park, 2005, p. 79) .

This whole range of intra-alliance contacts and negotiations can be subsumed under one multistage dynamic game. The bargaining issue can be defined as the stance that the Alliance intends to take jointly towards North Korea, and the bargaining pool

might be the total of benefits of achieving a peaceful solution of the North Korean nuclear crisis. Possible physical outcomes of the game were $t = (t_{Lib}, t_{Ukr})$, i.e. adopting either the Libyan model of dealing with the North (t_{Lib}) or the Ukrainian one (t_{Ukr}). Each of the two outcomes would possibly deliver the total of the following utilities

$$u_{US} = (u_{US_{Lib}}, u_{US_{Ukr}}),$$

$$u_{SK} = (u_{SK_{Lib}}, u_{SK_{Ukr}}),$$

$$u_{Jap} = (u_{Jap_{Lib}}, u_{Jap_{Ukr}}).$$

The game took the course of the US suggesting to its allies t_{Lib}, i.e. the Libyan model. It entailed the following utilities:

$$u = (u_{US_{Lib}}, u_{SK_{Lib}}, u_{Jap_{Lib}}).$$

South Korea's last offer was t_{Ukr}, with the utilities:

$$u = (u_{US_{Ukr}}, u_{SK_{Ukr}}, u_{Jap_{Ukr}}).$$

Japan's last offer was t_{Ukr}, too, leading to the same utilities.

If the US had accepted South Korea's last offer t_{Ukr}, then it would have obtained the utility payoff $u_{US_{ukr}}$ with certainty. On the other hand, if it insisted on its last offer t_{Lib}, it would either obtain the payoff $u_{US_{Lib}}$ (if South Korea and Japan accepted the offer) or would obtain the conflict payoff $c_{US_{Lib}}$. Having assigned the probability $(1-p)$ to the latter option, the US decided to insist on its preference and received the expected payoff:

$$(1-p)u_{US_{Lib}} + pc_{US_{Lib}} \geq u_{US_{Ukr}}$$

This was possible as the US did not assign particularly high value to the conflict payoff received after the disagreement with South Korea.

South Korea deployed a reasoning mirroring the one of the US receiving its expected payoff:

$$(1-p)u_{SK_{Ukr}} + pc_{SK_{Ukr}} \geq u_{SK_{Ukr}}.$$

As Japan's preferences matched the ones defined to picture the reasoning of the Americans, and none of the allies showed willingness to move ahead and make concessions, no alliance-level preference was reached.

Joint Statement of September 19, 2005 was the first tangible outcome of multilateral negotiations, agreed upon after the second phase of the fourth round of the Six-Party Talks. Therein, the allies found a middle ground, jointly declaring they had no nuclear weapons on the Korean Peninsula. Following its overall government philosophy of engagement of North Korea through inducements, South Korea gave a pledge to supply the North with 2 million of kilowatts of power (Wit, Wolfsthal & Oh, 2005, p. 77).

After a serious setback in negotiations and North Korea's announcement on October 9, 2006, that it had successfully tested plutonium-based weapons, differences in opinion among allies arose again. While the South Korean and Japanese governments were pressed by their own conservatives at home, they both decided to go for a more cautious reaction internationally and vis-à-vis the US. The allies worked together on drafting a UN Security Council resolution condemning North Korea's nuclear test and introducing a range of economic and military-trade related sanctions. And while both US allies asked from their senior partner to reconfirm its security commitment to them, South Korea rejected the idea of a preemptive strike threat. It also openly questioned the US hard-line approach focused on future sanctions and preconditions for any negotiations. Seoul feared this coercive attitude might provoke an aggressive response from Pyongyang. The allies were angry with each other: the Americans with South Korea for not going along, and the South Koreans were angry about the US ignoring all the diplomatic progress and gains South Korea had made (Park, 2005, pp. 80-86; Zissis & Lee, 2008).

Once the Action Plan of February 13, 2007 was concluded, restating North Korea's obligation to shut down and seal the Yongbyon nuclear facility, including the reprocessing facility, and invite back IAEA inspectors to conduct all the necessary monitoring and inspections, South Korean diplomats were greatly satisfied about the fact their approach to North Korea proved largely right. However, the reality of intra-alliance relations showed a great divergence of interests among allies. Japanese concerns were mostly attached to treatment of Japanese individuals abducted by the North Korean authorities in 1970s and 1980s. It resulted in the Japanese conditioning of any substantial

economic or political support to North Korea before the issue was settled (Zissis & Lee, 2008).

5.3 Completing the meta-model of relations within the Transpacific Alliance

Taking on the structure of the meta-model theoretically rounded up in chapter 4, it is possible now to finalize the first, C-dimension of the meta-model of intra-alliance relations.

1. *Alliance members –players*:

- United States, South Korea, Japan

2. *Fundamental alliance accord – legal game regulator*:

- Republic of Korea-US Mutual Agreement of 1954, 1948 UN Declaration of the Republic of Korea; Japanese Constitution of 1946

3. *Power flow and capabilities distribution within alliance – pragmatic game regulator*:

- Capabilities distribution ratio within the alliance: US: SK: Japan = 8:1.7:0.3

4. *Allies' common and compatible interests*:

- Identical strategic (long term interests) and compatible short and mid-term interests

5. *Intra-alliance cooperation as coordination and mutual accommodation –*

- As demonstrated in the game induced by the North Korean nuclear program, the frequency of the preference accommodation procedure deployment is slightly higher than the ordinary preference aggregation, that, in the intra-alliance sequential bargaining without a voting, i.e. a transformation principle leads to inherent conflict payoffs, or disagreement values.

The DF-dimension of the meta-model is as follows.

1. Intra-alliance stimuli:

- *Allies' varying interests and motivation to ally with other actors:*

- Given the security environment, the Transpacific Alliance holds firmly together with, nevertheless, unevenly distributed interest in its maintenance (stronger with South Korea and Japan)

- Allies' particular definitions of *national interest*:

- Highly compatible strategically, often conflicting tactically

- *Individual member's internal politics, socio-economic, demographic and military conditions, political and regime similarity of allies:*

- The game depicted in chapter 5 showed a highly obstructive influence of domestic politics on intra-alliance bargaining, regularly stiffening allies' positions and preferences. Lack of regime similarity, however, proved to be less relevant a factor, in contrast to socio-economic interdependence of South Korea and Japan on the US.

- *Relative significance of issues with converging/diverging interests among allies, resulting parallelogram of allies' diverse converging/diverging interests* :

- A 'diamond structure' of the parallelogram: short term interests often diverging as opposed to long term interests with a pronounced convergence.

- *Allies' relative capabilities*:

- US: South Korea: Japan = 8:1.7:0.3

- Allies' strategies of intra-alliance behavior:

- As seen in the North Korean game, buck-passing and free-riding often deployed by South Korea and Japan; chainganging occasionally by Japan and bandwagoning never. Intra-alliance tit-for-tat rarely deployed; in a limited number of cases by South Korea due to publicity demands of its domestic constituency

- *Intensity, extensity and density of information flow among allies*:

- Very high and on a broad variety of issues, helped reconcile intra-alliance differences in a number of cases

- *Allies' individual and common historical and security legacy*:

- Extremely intertwined, approximately to the level of short-term security and survival dependence of South Korea and Japan on the US. Simultaneously, the US depends on its Northeast Asian allies to deter North Korea, as they increase the autonomy of the American military by boosting its power projection perspectives.

2. *Extra-alliance (environment's) stimuli*:

- High level of influence of North Korea’s behavior on the state and dynamics of intra-alliance relations, as proved in the game developed in this chapter. It is demonstrated that the Transpacific Alliance, its life and survival greatly depend on actions, bargaining strategies and tactics of North Korea, especially as its regime poses a supreme security threat for the Allies. Moreover, North Korea proved to be the key factor of relations within the Transpacific Alliance.

6. CONCLUSIONS AND FURTHER REASEARCH RECOMMENDATIONS

Aiming to shed some more light on a methodological dilemma related to the applicability problem of quantitative methods of analysis to diverse phenomena of international affairs, opened this study was incepted by asking how effectively an intersection of game theory and theories of alliances can explain dynamics of intra-alliance relations, causes and outcomes of allies' behavior, motives and interests that trigger and drive that behavior. To answer such a theoretically comprehensive question, several steps have been taken in the course of the research. First, it was stated that answering this question demands creating an inclusive model of analysis that can be later applied to specific cases. We decided to call this model a "meta-model" as it encompasses two large theoretical inputs, each of them being a particular model of analysis itself. Then, the meta-model was being built up by initially framing input1 out of the first theoretical field we concentrated on, the field of theories of alliances. After framing this first input and shaping a unique logically and substantially coherent structure out of specific theories (chapter 2), adjusted the analytical apparatus of game theory was adjusted to substantial and philosophical demands of theories of alliances and to the nature of intra-alliance relations more generally, framing input2 to the meta-model. Closing the meta-theoretical part of the study, these two inputs were joined, ending in a construction of the final meta-model of analysis of intra-alliance relations. In order to practically observe the explanatory reach of the meta-model, it was finally applied to the empirical case of intra-alliance relations between the US, South Korea and Japan, who form the so-called Transpacific Alliance.

After such an abundant vocation of examining the methodological and meta-theoretical quandaries of International Relations and security studies, the following conclusions on the applicability of the meta-model are reached. Furthermore, research hypotheses presented in the opening chapter are confirmed, too.

Firstly, referring to the hypothesis $_{G}MTH_{1}$, it was proved that game theory's tools are highly applicable to the study of intra-alliance relations, as its five defining elements (players, strategies, information distribution, payoff consequences and preferences over payoffs) are entirely applicable to the inquiry of components and determinants of

international politico-security alliances. A strong correlation between these two sets of elements is confirmed through theoretical tests presented throughout the chapters 2 and 3, and in the course of application and empirical testing of the meta-model. Components and determinants of the Transpacific Alliance enlisted in chapter 5 are either derived directly from (e.g. intra-alliance information exchange and distribution) or confirmed by the game theoretic analysis in chapter 4 (e.g. the factor of extra-systemic stimuli).

Secondly, concerning the hypothesis $_{G}MTH_{2}$, the application of the meta-model in chapter 4 has demonstrated that allies tend to formulate their preferences individually factoring in an independently perceived collective (alliance-level) interest.

Thirdly, in reference to the general hypothesis $_{G}MTH_{3}$ and specific hypothesis $_{S}MTH_{1}$, this study has confirmed that international politico-security alliances exert an independent influence on the formation of allies' preferences, yet predominantly at the individual level of alliance-members. Along with the compelling theoretical arguments presented in chapter 2 and implemented in chapter 3, the case of relations among the US, South Korea and Japan has proved that allies tend to take most important decisions following an intra-alliance preference accommodation procedure, the proof of a discernable impact of their joint arrangements such as international politico-security alliances. It means that in spite of their particularistic concerns, allies tend to mutually adapt their preferences when it really matters, as it was shown in the case of the Agreed Framework of 1994. They do it as they are compelled by the evolving process of intra-alliance cooperation and coordination. Their mutual commitments, so often recalled in the case of relations among the US, South Korea and Japan, tie them together and make them depart from exclusively individual gains, often suffering short-term losses. International alliances can indeed, through complex interactions of players produce independent effects on international affairs. They are institutions of international politics, more then a simple aggregate of their members (proved by incorporating the preference accommodation procedure into the meta-model).

Regarding the second specific hypothesis $_{S}MTH_{2}$, the construction of input2 and the empirical tests the meta-model was put to proved that alliance-level preferences cannot be simply disaggregated into their individual preference-components, as it is

particularly with the case of the South Korean and Japanese preferences on the Agreed Framework of 1994.

As for the *empirical* strand of the research, the most general hypothesis $_{G}EH_{1}$ that the overall tendency in negotiating strategies of the Transpacific Alliance is to the one of ambiguity is confirmed as well. It has been demonstrated that the Alliance mostly maintains an overtly tough and firm posture towards North Korea accompanied by a harsh public discourse, whereas in practice it continues with the policy of intense efforts to involve and commit North Korea to talks with International Atomic Energy Agency and South Korea. Secondly, this ambiguity is an alliance-level preference reached through a preference accommodation process (hypothesis $_{G}EH_{2}$).

This closure will also offer a set of conclusions that are not directly related to the research hypotheses enlisted above.

Firstly, game theory does recognize the impetus for cooperation in international politico-security alliances, in spite of its largely competitive epistemological setting, as it allows for a considerable deployment of its rather collaborative versions (coordination game, El Farol Bar problem or minority game) in explaining dynamics of intra-alliance relations.

Secondly, game theory convincingly demonstrates how information exchange among allies increases prospects for Pareto optimal solutions of games with a Nash equilibrium, thus proving its critical importance for the state and development of intra-alliance relations as far as actors involved with an alliance see that engagement viable.

However, some limitations to applicability of game theory to the study of intra-alliance relations can be identified as well, even after a massive adaptation of its tools is done in chapter 3.

Firstly, game theory can host continuing logical fallacies. As demonstrated in our case study (chapter 5), allies can play their most rational and reasonably defined most preferred strategies (with both individual and alliance-level interests calculated) and still end up with suboptimal game solutions that do not fully (or at all) satisfy their individual nor collective preferences, as in the case of the South Korean tacit rejection of the February 13 Action Plan.

Secondly, game theory can hardly simultaneously combine allies':

1. individual preferences,
2. collective preferences,
3. cooperative decision-making through preference accommodation, and
4. competitive game playing through a simple preference aggregation.

As seen in the case study (empirical section of the study), procedures of game framing impose a rather fragmented deployment of these four elements, which, in turn, poses a severe limitation on describing and explaining the course of intra-alliance relations. However, practical solutions of such quandaries are found through moderate simplifications and game complexity rationalizations, as done in the case study by dividing the game in two somewhat separated periods.

Thirdly, game theory can easily be manipulated to a level of false or ostensible rationality. Although it aims at taking an exact and rational approach to phenomena of international politics, it remains impotent in defining what is truly and intrinsically rational. We can, using the example of our case study, plausibly claim that taking an uncoordinated aggressive approach towards North Korea by its northern neighbor was at times far from rational, obstructing the negotiation process between Washington and Pyongyang. This is a somewhat general deficiency of game theory, as it does not manage to successfully combine insights of procedural and substantial rationality (Isla, 2000, pp. 347-363), mostly relying on the former. If allies' strategies, payoffs and preferences are correctly defined in terms of game theory and its formal rules, then they are qualified as rational. This is, nevertheless, only the procedural level of rationality, not the substantial one. Still, even if we want to frame a model that would encompass both types of rationality, the problem of operationalizing substantial rationality would remain. What is substantially rational may depend on subjective personal preferences, cultural and social background of players or researches, their particular interests related to the object of study. The well-known problem of cognitive dissonance finely encapsulates this quandary, showing how personal mindset can frame almost anything as rational. All these deficiencies leave a large room for manipulations with game theory by both researches and policymakers, and this is where reasons for the failure to mutually adapt 'rational' strategies and 'irrational' outcomes should be looked for.

Once the lessons on the applicability of game theory to the study of intra-alliance relations are drawn, some research recommendations on how to cope with the limitations and deficiencies in its application should be given, too.

Firstly, it is greatly fruitful to examine data on empirical events (primary and secondary sources) as thoroughly as possible in order to reach a satisfactory level of certainty when a game is framed (especially strategies that players have available and their expected payoffs). This would considerably marginalize speculations on such qualifications as 'rational' and 'irrational', as we would at least know the entirety of a specific player's position and relevant circumstances that act as determinants of its decision-making.

Secondly, researches and policymakers should seriously examine the dataset they have available when rationalizing the empirical events down to only five game elements. This process of qualification and elements identification poses serious traps when simplifications and abstractions are being made. That is why the simplifications that are always needed and desirable must not lead to inadequate game framing, where five game elements do not correspond to empirical events. This takes us back to the notion of scientific relevance and representativeness and how it is supposed to be deployed practically (Tenenbaum & Griffiths, 2001). Although certain criteria are set (e.g. scope of topics covered in a dataset, extent of inclusion of background variables, relation between the former and the latter, sufficiency of the sample etc.), researchers and policymakers are left on their own when it comes to actually applying them, as various and distinct assessments are repeatedly made.

Thirdly, analytically decomposing and dissecting a game can be fruitful if done in a controlled way and if a re-inclusion and re-composition of all the analytically derived components is added to the process. It means that a dialectically dynamic method must be preserved if the study object is to be regarded in its entirety. As an example of this decomposition, we can use the partition of the game that we did in our case study (by dividing it in two periods with two distinct decision-making procedures) in order to discern individual and collective preferences, preference aggregation and accommodation.

Finally, getting back to our initial question on the practical reach of the intersection of game theory and theories of alliances, we believe that the arguments presented prove this intersection can be highly fruitful, under the strict condition of logically and substantially coherent combination. Moreover, we remain aware of the abstract and meta-theoretical domain and scope of this study, which is why we believe more operationalizing studies (recommended at several points throughout the paper) with a stronger empirical orientation are needed to complement it. However, to conduct such examinations, a comprehensive model of study of intra-alliance relations is needed. Consequently, providing researches and policymakers with one was the dominant and driving intention of this paper.

BIBLIOGRAPHY:

Allan, P. & Schmidt, C. (Eds.). (1994). *Game Theory and International Relations*. *Preferences, Information and Empirical Evidence.* Cambridge: Edward Elgar Publishing Limited

Altfeld, M. F. (1984). The Decision to Ally: A Theory and Test. *Western Political Quarterly*, 37, pp. 523-544

Arrow, K. J. (1963). Uncertainty and the Welfare Economics of Medical Care. *American Economic Review*, 53 (5): pp. 941-73

Bacharach, M. (1993). Variable universe games, in *Frontiers of Game Theory*, pp. 255-75.

Bacharach, M. (1999). Interactive team reasoning: a contribution to the theory of cooperation. *Research in Economics*, 53:117-47

Bacharach, M. & Lawler, E. J. (1981). *Bargaining: Power, tactics, and outcomes*, San Francisco, CA: Jossey-Bass

Bearce, D. H., Flanagan, K. M. & Floros, K. M. (2006). Alliances, Internal Information, and Military Conflict among Member-States. *International Organization*, 60, pp. 595-625

Bennett, D. S. (1997). "Testing Alternative Models of Alliance Duration, 1816-1984." *American Journal of Political Science* 41 (July):846-878.

Brams, J. S. (2000). Game Theory: Pitfalls and Opportunities in Applying It to International Relations. *International Studies Perspectives*, pp. 221-232

Brams, S. J. (2001). Theory of Moves: Game Theory and the Cuban Missile Crisis. *Plus Magazine*. University of Cambridge

Bush, G.W. (2003). *President Bush Addresses the Nation*. The White House. Retrieved May 20, 2007, from http://www.whitehouse.gov/news/releases/2003/03/20030319-17.html

Buzan, B., Weaver, O. & De Wilde, J. (1998). *Security: A New Framework for Analysis*. London: Lynne Rienner Publisher

Christensen, T.J. & Snyder, J. (1990). Chain Gangs and Passed Bucks; Predicting Alliance Patterns in Multipolarity. *International Organization*, 44, 2

Clemen, R. (1996). *Making Hard Decisions* (2nd edition). Belmont, MA: Duxbury Press

Cordesman, A. H. & Burke, A. A. (2002). *Western Military Balance and Defense Efforts. A Comparative Study of Military Expenditures; Manpower; Land, Air, Naval, and Nuclear Forces*. Washington DC: Centre for Strategic and International Studies

Croft, G. W. (1996). *Glossary of Systems Theory and Practice for the Applied Behavioral Sciences*, Syntropy Incorporated, Freeland, WA, Prepublication Review Copy

Cronin, R.P. et al. (1994). *North Korea's Nuclear Weapons Program: U.S. Policy Options*. CRS Report for Congress

Deutsch, K. et al. (1957). *Political Community in the North Atlantic Area.* Princeton: Princeton University Press

Downs, C. (1999). *Over the Line: North Korea's Negotiating Strategy*. Washington D.C.: The AEI Press

Galitzi, C. (1933). The Balkan Federation. *The Annals of the American Academy of Political and Social Science,* 168, 1, pp. 178-182

Gauthier, D. (1975). Coordination. Dialogue, 14:195-221

Gold, N. & Sugden, R. (Eds.). (2006). *Beyond Individual Choice. Teams and Frames in Game Theory*. Princeton, NJ: Princeton University Press

Guillermo, O. (2001). *Game theory* (3rd edition). San Diego: Harcourt, Inc

Gerth, H. & Mills, W. (1946). *Science as Vocation: From Max Weber: Essay in Sociology*. New York: Oxford University Press

Gilinsky, V. (1997). *Nuclear Blackmail: The 1994 US-Democratic People's Republic of Korea Agreed Framework on North Korea's Nuclear Program*. Stanford: Hoover Institution Press

Henriksen, T.H. & Mo, J. (1997). *North Korea after Kim Il Sung: Continuity or Change?* Stanford: Hoover Institution Press

Hollis, M. (1998). *Trust within Reason*. Cambridge University Press

Hodgson, D. H. (1967). *Consequences of Utilitarianism*. Clarendon Press

Isla, A. (2000). From procedural to complex rationality relations: Observed system and observing system. *European Journal of Economic and Social Systems*, 14, 4

Istvan Deak. (1979). *The Lawful Revolution, Louis Kossuth and the Hungarians, 1848-1849* . Columbia University Press: New York

Katz, A. & Butler, B. (1994). *"Game Commander"-Applying an Architecture of Game Theory and Tree Lookahead to the Command and Control Process*. Proceedings of the Fifth Annual Conference on AI, Simulation, and Planning (AIS94), Florida

Kennedy, P. (1987). *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000.* New York: Random House, Inc.

Keohane, R. O. & Nye, J.S. (2000). *Power and Interdependence. Third Edition.* New York: Pearson Longman

Keohane, R. O. (1984). *After Hegemony: Cooperation and Discord in the World Economy*. Princeton, NJ: Princeton University Press

Keohane, R. O. (1998). International Institutions: Can Interdependence Work? *Foreign Policy*, 110

Keohane, R. O. (1989). *International Institutions and State Power*. Boulder, Colo.: Westview Press.

Kimball, A. L. (2006). Alliance Formation and Conflict Initiation: The Missing Link. *Journal of Peace Research,* 43(4), pp. 371-389.

Kim, W. (1991). Alliance Transitions and Great Power War. *American Journal of Political Science*, 35, 4, pp. 833-850

Krasner, Stephen D. (1983). Structural Causes and Regime Consequences: Regimes as Intervening Variables. *International Regimes,*. 1-21. Ithaca, New York: Cornell University Press

Kyuong-Soo, K. (Ed.). (2004). *North Korea's Weapons of Mass Destruction: Problems and Prospects*. Elizabeth, NJ: Hollym

Lai, B. & Reiter, D. (2000). Democracy, Political Similarity and International Alliances, 1816-1992. *Journal of Conflict Resolution*, 44, 203

Leeds, B. A. (2003). Alliance Reliability in Times of War. *International Organization*, 57, pp. 801-827

Morrow, J. D. (1991). Alliances and Asymmetry: An Alternative to the Capability Aggregation Model of Alliances. *American Journal of Political Science*, 35, 4, pp. 904-933

Niksch, L.A. (2004). *Korea: US-Korean Relations*. CRS Issue Brief for Congress

Oberdorfer, D. (1997). *The Two Koreas: A Contemporary History*. Reading, Massachusetts: Addisson-Wesley

O'Hanlon, M. & Mochizuki, M. (2003). *Crisis on the Korean Peninsula: How to Deal with a Nuclear North Korea*. New York: Brookings Institution

Park, J.S. (2005). Inside Multilateralism: The Six-Party Talks. *The Washington Quarterly*, 28:4, pp. 75-91

Regan, D. (1980). *Utilitarianism and Cooperation*. Clarendon Press

Renzi, W. A. (1968). Italy's Neutrality and Entrance into the Great War: A Re-examination. *The American Historical Review*, Vol. 73, No. 5, pp. 1414-1432

Salmon, T., (2005), "The European Security and Defence Policy: Built on rocks or sand?" *European Foreign Affairs Review*, Vol. 10, 359 – 379

Sigal, L.V. (1998). *Disarming Strangers: Nuclear Diplomacy with North Korea*. Princeton University Press

Smith, A. (1995). Alliance Formation and War. *International Studies Quarterly*, 39, pp. 405-425

Stockholm International Peace Research Institute (2001). *The SIPRI Military Expenditure Database*. Retrieved May 5, 2007, from http://first.sipri.org/non_first/milex.php

Sugden, R. (2000). Team Preferences. *Economics and Philosophy*, 16, pp.175-204

Taliaferro, J. (2000). Security Seeking Under Anarchy, Defensive Realism Revisited. *International Security*, 25, 3, pp. 128-161

Tenenbaum, J. B. & Griffiths T. L. (2001). The Rational Basis of Representativeness. *23rd Annual Conference of the Cognitive Science Societ*y, pp.1036-1041

Trevor, S. (2005). European Union Security and Defense Policy: Built on Rocks or Sand? *European Foreign Affairs Review*, 10, pp. 359-379

Trivanovitch, V. (1931). Serbia, Russia and Austria During the Rule of Milan Obrenovich, 1868-1878. *The Journal of Modern History*, 3, 3, pp. 411-440

Tuomela, R. (1995). *The Importance of Us*. Stanford University Press

US-North Korean Joint Statement. June 11, 1993. Geneva

Waltz, K. N. (1979). *Theory of International Politics*. New York: McGraw-Hill Publishing Co.

Wallner, K. (2002). The Provision of Public Goods in International Relations: A Comment on "Goods, Games, and Institutions". *International Politics Science Review*, Vol. 23, No. 4, pp. 393-401

Williams, P., Goldstein, D. M. & Shafritz, J. M. (2006). *Classic Readings and Contemporary Debates in International Relations*. Belmont, CA: Thomson Wadsworth

Williams, M. C. (2003). Words, Images, Enemies: Securitization and International Politics. *International Studies Quarterly* 47, pp. 511-531.

Winston, P. (1992). *Artificial Intelligence*. Reading Massachusetts: Addison-Wesley

Wit, J.S., Poneman, D.B. & Gallucci, R.L (2004). *Going Critical: The First North Korean Nuclear Crisis*. Washington, D.C.: The Brookings Institution

Wit, J.S., Wolfsthal, J. & Oh, C.S. (2005). *The Six Party talks and Beyond*: *Cooperative Threat Reduction and North Korea*. Washington, D.C.: Center for Strategic and International Studies

Young, J. & Kent, J. (2004). *International Relations since 1945. A Global History*. Oxford: Oxford University Press.

Zissis, C. & Lee, Y. (2008). *The US-South Korea Alliance*. Council on Foreign Relations. Retrieved June 2, 2008, from http://www.cfr.org/publication/11459/ussouth_korea_alliance.html