

Confidence analysis for nuclear arms control: SMT abstractions of Game Theoretic Models

Paul Beaumont¹, Neil Evans², Michael Huth¹, and Tom Plant²

¹ Department of Computing, Imperial College London, London, SW7 2AZ, UK
{paul.beaumont, m.huth}@imperial.ac.uk

² AWE Aldermaston, Reading, Berkshire, RG7 4PR, UK
{Neil.Evans, Tom.Plant}@awe.co.uk

Abstract. We consider the use of game theory in an arms control inspection planning scenario. Specifically we develop a case study that games the number of inspections available against an ideal treaty length. Normal game theoretic techniques struggle to justify pay-off values to use for certain events, limiting the usefulness of such techniques. In order to improve the value of using game theory for decision making, we introduce a methodology for under-specifying the game theoretic models through a mixture of regression techniques and Satisfiability Modulo Theory (SMT) constraint solving programs. Our approach allows a user to under-specify pay-offs in games, and to check, in a manner akin to robust optimisation, for how such under-specifications affect the ‘solution’ of a game. We analyse the Nash equilibria and the mixed strategy sets that would lead to such equilibria - and explore how to maximise expected pay-offs and use of individual pure strategies for all possible values of an under-specification. Through this approach, we gain an insight into how - irrespective of uncertainty - we can still compute with game theoretic models, and present the types and kinds of analysis we can run that benefit from this uncertainty.

1 Introduction

Arms control agreements - and the verification processes that are often dictated by them - are important tools that help nations manage their security relationships with each other. It is possible for agreements to operate without a verification mechanism. Indeed the design and implementation of these processes can be complicated, intrusive and expensive - but the absence of such a deterrent to cheaters clearly limits the potential effectiveness of agreements. It is normally accepted that some level of inspection regime is a necessary element of any such regime.

Research into the effectiveness of different technologies and procedures to support nuclear arms control inspections is being pursued increasingly widely - the UK-Norway Initiative [19], ongoing UK/US collaboration [14], and the International Partnership for Nuclear Disarmament Verification (IPNDV) [15] are a few examples of contemporary efforts. Reciprocity is often a key principle of these processes; in this circumstance a nation must balance the potential added value from a particular technique against any potential detriment to its national security interests from a reciprocal inspection utilising the same technique.

All of these tensions - cost, value added, reciprocity - motivate the need for tools that model, analyse and optimise verification processes in this domain. Given that the stakes for arms control decisions are particularly high and the circumstances in which they are made can be very complex, we are additionally motivated to examine ways in which those decisions can be supported by robust and algorithmic analytical tools. If the output of these tools can be in some sense provably correct - which might be difficult to check analytically for all but the most trivial cases - then that would be an additional valuable benefit.

Game-theoretic frameworks were heavily developed for general arms control, and have been employed in the past to analyse aspects of the nuclear problem space. For example, see the normal form games of [20, 21]. Their main limitation is that conventionally, a modeller must attach pay-off values to events, which in practice can be difficult to justify. Bayesian Belief Networks (BBNs) have been used to consider particular inspection scenarios [5, 6], and to see how the uncertainty inherent in them can be modelled as their use in probabilistic reasoning can incorporate subjective as well as objective evidence or beliefs - which work in [5, 6] sought to justify. Meanwhile, the work in [7] uses dynamical

systems to model the uncertain, temporal interactions and human decision making of parties subject to an arms control process.

In this case study, we have two parties that are negotiating the length of a treaty and the number of inspections that will be held during it. In our model this reduces to a consideration of the number of inspections per unit time. We focus on how to best choose an apt maximum number of inspections allowed for a treaty of an undetermined length. Too few inspections leads to increased risks of treaty collapse through insufficient deterrent to cheating or insufficient confidence that the other party is complying with its obligations; too many inspections and the cost of verification becomes too much to bear. We thus model the “*number of inspections:treaty length*” ratio in a game-theoretic manner. The pay-offs for the games will be based on those used in an inspection game introduced by Dresner [10], discussed further in Section 4, that addresses the issue of how many inspections to run in a fixed length treaty. We will allow these pay-offs to be underspecified and show how we can thus reason about a whole set of games with identical structure so that any insights gained (e.g. Nash equilibria) will be robust under changes of pay-off values within specified uncertainties. We hope that this will illustrate how our approach can improve the methodological value of game theory for this application domain.

2 An overview of game theory

Game theory - the ‘study of mathematical models of conflict and cooperation between intelligent, rational decision makers’ [13] - has an established history in defence analysis, perhaps most famously in the context of nuclear deterrence (see for example [18]). In an arms control context, an inspecting party seeks data to support an assessment of another party’s compliance with an agreed set of rules. Whilst both parties have an interest in being seen to comply with this set of rules, other national interests may differ. This in turn leads to different objectives for the parties involved, the pursuit of which can be analysed by studying mathematical representations of inspections in a non-cooperative two-player game.

Games are generally characterised by whether they have zero-sum or non-zero-sum pay-offs to players, and whether they are non-cooperative or cooperative (in the latter case, coalitions of players against others may form). Games are often given in ‘normal form’: as an array of matrices (representing an N-dimensional matrix) storing the pay-offs to players for the ‘pure’ strategies of each of the N-players (as in Figure 1). ‘Extensive form’ games represent the sequence of players’ moves and decisions in a tree-like graph (as in Figure 2).

In reality, the players will more often than not have to choose ‘mixed strategies’: probabilistic combinations of their pure strategies - to achieve optimum results. This probability distribution on the set of pure strategies for a player therefore becomes for us a focus of interest as we seek the (possibly multiple) equilibrium solutions to games. Other variables of interest might include the expected pay-offs for each player, and the stability or otherwise of the equilibria to changes in the parameters of the game. These may either be derived for the equilibrium or, in the case of the latter, may need to be calculated separately in the model.

We recall key concepts of game theory (rather informally) below:

Normal Form Games written in a pay-off matrix, with the strategies of each player denoted along the rows, columns (or in cases of more than 2 players, other N dimensions of the matrix). See Figure 1.

Extensive Form A game specification that shows the decisions available to players based on other players’ moves, or moves at random, and their associated pay-offs. See Figure 2.

Pure strategy A complete description of the move made in a game (played once, or played many times) and is a distinct strategy.

Mixed strategy A probabilistic mixture between multiple pure strategies, in the ratio of which these are executed at random when the game is played repeatedly.

Value of a game The expected average pay-off to the player should infinitely many copies of the game be played. Note that there is a subtle difference between studying multiple independent

runs of a game with no player learning between games, and studying repeated games with player learning. The value of a game is derived from an analysis of the former case.

Solution A mixed strategy for each player that forms a ‘set’ of strategies for the respective player. Many different choices (or ‘best’ choices), of which the Nash equilibrium is one.

Nash equilibrium The ‘*solution*’ of a game in which no player can achieve a higher value by changing their strategy while other player strategies remain unchanged, and therefore no player has an incentive to change strategy.

When we allow games to contain symbolic rather than numeric information it is not immediately clear how the above notions all generalise, or can be computed. The main issue arises around how to handle under-specifications with respect to what we refer to here as ‘boundary points’. The boundary points are the numerical or symbolic values that affect players’ decision making, and therefore computation of the Nash equilibrium and value of a game. All such computations are based on taking the maximum or minimum pay-offs in a row or column. This means the relative value of the pay-offs is important and affects the result. Taking the case of the game represented in Figure 1, the points at which, for instance, a maximisation function $\max(a_{i_1j_1}, a_{i_1j_2}, a_{i_1j_3}, x)$ changes values (as x varies) are our important boundary points.

We can explain the issue of boundary points as analogous to the situation of having a Control Flow Graph that switches cases as x lies in between other pay-off values. Any tool that offered a solution would have to respect these boundaries and present results for symbolic under-specifications that spanned ‘boundaries’. There is value in under-specifying games to symbolic computations though. Solving such games - often known as ‘parametrised games’ - can be done analytically, by considering the cases in which each of the pay-offs interact on a boundary point. This approach is not efficient, however, and becomes cumbersome for games of increasing sizes (for instance, for many strategies, or many players). We seek a tool that can automate the “logical switches” associated with the above boundary cases and can answer queries about such games of interest to inspection regimes.

$A =$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & x & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Fig. 1. For pay-off matrix A , the value of pay-off x relative to pay-offs a_{12}, a_{32}, a_{21} and a_{23} (those in the same rows or columns) will determine the equilibria of the game (i.e. for the columns, if $a_{12} \leq x \leq a_{32}$ or if $a_{12}, a_{32} \leq x$ etc) . In our approach, we would like to, and we will determine the equilibria for all such cases.

The capabilities that we developed for our approach are captured in a Python package that can handle N -players with any finite number of strategies for each player, and multiple under-specifications of the pay-off values, with different symbolic values. Our tool has the ability to compute all Nash equilibria of a game (or some, if only some are required), the ability to compute (mixed) strategies that realise such equilibria, the expected pay-offs to players, and the ability to determine how an under-specification impacts the sensitivity of a game’s solutions.

3 Implementation

There are currently two widely used game theory software tools, *The Gambit Project* (known as *Gambit*) [11] and *Sage* [17].

Gambit is designed for N player non-cooperative games, and can handle neither cooperative games, nor symbolic games. It can calculate Nash equilibria, and the strategies required to build those equilibria. *Sage* implements *Gambit* for 2-player games only (and uses this to power its solution engine for non-cooperative normal form games). It additionally includes characteristic form games and implements code to find the solution of these when cooperative games are presented in characteristic

function form. *Sage* provides no means to go from cooperative normal form to characteristic function form. Neither tool supports symbolic (parametric) games.

We automate this ‘symbolic solving’ part of this work by computing Nash equilibrium results numerically (using *Gambit*) at intervals $l + t \times \frac{u-l}{10}$ across our symbolic area of interest $l \leq x \leq u$, (for t in $\{0, 1, 2, \dots, 10\}$) and then using the computed solutions to come up with a (regression) formula that faithfully represents the numerical results from these intervals in an approximated symbolic manner. We decide to accept a regression when the r^2 value hits a modeller’s acceptable threshold (a regression metric indicating a good fit), and bisect and split the problem over any range that doesn’t meet this threshold (which normally coincides with crossing a boundary point). This will partition intervals into units of possibly varying length that all provide a good enough fit. Once a regression formula is found for the whole under-specified interval, it is asserted into the SMT Solver as a mathematical statement, as described below.

Satisfiability modulo theories (SMT) [3,12] is an approach to automated deduction that checks the satisfiability of logical statements. We choose Z3 [3] as our SMT solver, though it would be relatively easy to replace it with another solver such as CVC3 [4]. Z3 reasons with mathematical statements that we assert as constraints and seeks to find satisfiability witnesses for cases where all constraints are met. For instance, we could define a variable x and say $5x^2 + 6x + 1 = 0$, and Z3 might return $x = -1$, which is a solution to the equation. We could additionally say $x \neq -1$, and Z3 would return $x = -0.2$; but stating that that isn’t true either would make Z3 determine that the current system of constraints is not satisfiable.

We utilise *Gambit*’s Python API to interact with the N -player Nash equilibrium solver engine. The rest of our code is in pure Python, utilising the Z3 SMT Solver API to handle symbolic decision making, and we have made these available as two Python classes, which we have called: `SymbolicGames` and `ExtensiveGames`. `ExtensiveGames` is an implementation of the Dresher models. `SymbolicGames` (which will deal with non-cooperative games in general - for any $N > 1$ player, normal form games) is concerned with the regression procedure itself and answers the queries posed in this paper.

We move on to showcase how our methodologies, as described, can then be employed by and benefit those who advise decision makers in an arms control scenario case study.

4 Scenario of interest

The extensive form ‘Dresher game’, $I(n, m)$, [1, 2], and seen in Figure 2, is a model that sought to assess the expected utility of having m inspections over a n time step period.

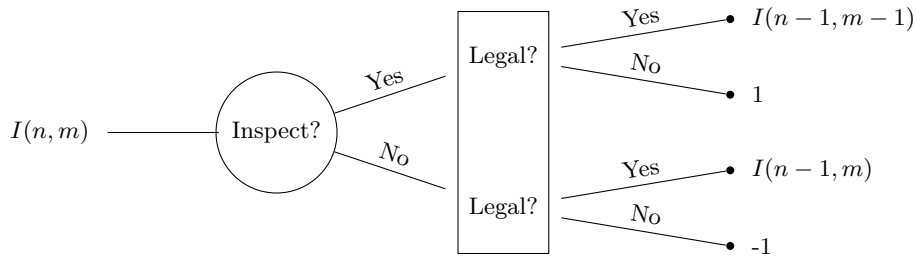


Fig. 2. Extensive form game of how to best schedule inspections. This is based on the model of Dresher and assigns payoffs based on whether an inspection was used wisely to detect cheating, or not. The modeller decides whether to inspect or not based on the maximal utility of the system, and the the legality of the inspected nation’s behaviour is left to chance (the probability of which can be the same throughout the model, set explicitly for different values of n and m , or can be assigned different probabilities randomly [1, 2]).

In these models, if we optimise our inspection decision for greatest reward (circular node), and take expectations over the legality of how the other nation state acts (rectangular node), then the output from this model is a maximum utility over treaty length n , for a certain initial number of

allowed inspection, m . If we vary n and m independently we generate multiple strategies that could potentially be incorporated into a treaty, and through which to choose from - or, potentially, for which scenarios to plan for.

The utility from this extensive game could then be processed as a pay-off and played as a normal form game to identify what are the best strategy options for different objectives. Running this model for varying probabilities of legality at each inspection stage, we generate a normal form example game, A (Figure 3), where the pay-offs correspond to the expected utility for an inspecting nation for varying values of n (rows) and m (columns). We under-specify the perceived pay-off when the inspector runs out of inspections and $m > n$ with $-x$ for $0.5 \leq x \leq 0.9$ to model that the inspector would likely want to conduct more inspections and so feel an overall detriment, but that the magnitude of this detriment is not readily specified, but resides in the interval $[0.5, 0.9]$.

$$A = \begin{bmatrix} \frac{31}{50} & \frac{41}{50} & -x & -x & -x & -x & -x & -x \\ \frac{229}{500} & \frac{329}{500} & \frac{419}{500} & -x & -x & -x & -x & -x \\ \frac{1561}{5000} & \frac{2561}{5000} & \frac{3461}{5000} & \frac{4271}{5000} & -x & -x & -x & -x \\ \frac{9049}{50000} & \frac{19049}{50000} & \frac{28049}{50000} & \frac{36149}{50000} & \frac{43439}{50000} & -x & -x & -x \\ \frac{31441}{500000} & \frac{131441}{500000} & \frac{221441}{500000} & \frac{302441}{500000} & \frac{375341}{500000} & \frac{440951}{500000} & -x & -x \\ -217031 & 782969 & 1682969 & 2492969 & 3221969 & 3878069 & 4468559 & -x \\ \frac{-6953279}{50000000} & \frac{3046721}{50000000} & \frac{12046721}{50000000} & \frac{20146721}{50000000} & \frac{27436721}{50000000} & \frac{33997721}{50000000} & \frac{39902621}{50000000} & \frac{45217031}{50000000} \end{bmatrix}$$

Fig. 3. Pay-off matrix for a player inspecting a nation state for varying treaty length strategies (rows - conventionally the second player's strategies - representing treaty lengths 2 to 8) and varying inspections (columns - the first player's strategies - representing inspection numbers 1 to 8). We are unsure over what probabilities to assign to events where $m > n$ without knowing if the inspecting nation would *wish* to hold more inspections. For simplicity, we model this as the same pay-off, x , throughout, though this could be extended in future work.

Negating this pay-off matrix and replacing under-specification x with z , for $0.5 \leq z \leq 0.9$, we create a pay-off matrix for the host nation (to add a second under-specification and form a non-zero-sum game). We call this pay-off matrix B . The game is 'zero-sum' for valid inspection routines, $m < n$, but for under-specifications $-x < 0$ in A (and under-specifications $-z < 0$ respectively in B), reflects that no nation wins (or potentially their trust is reduced) if they are unable to carry out any more inspections: it gives no opportunity to check or prove compliance, and so is assumed to be *definitely* a negative pay-off to both. The game is therefore non-zero-sum overall.

We noted above that reciprocity is a key feature of arms control processes. Thus, in our game the parties swap roles at various points; whilst 'players' in a normal form game normally would refer to different parties, here we use the game-theoretic approach as a way to solve a synthesis problem of how to choose best strategies for the host and inspectors. We refer to player A as the inspecting party using pay-off matrix A , and player B as the inspected (host) nation using matrix B . A and B are thus distinct from the nations party to the treaty. Each player has pay-off matrices as described above with under-specifications x and z respectively, and it is in the best interests of both nations that A and B are satisfied with the result of the game (and achieve a Nash equilibrium).

For this normal form game, there are questions whose answers would aid in decision support:

- 1) Given uncertainty x and z in the Dresher model over pay-offs to A and B, maximise the use of 1 inspection only and determine the values of x and z for which this would hold true. What other strategies involving 1 inspection form a Nash equilibrium?
- 2) For which x and z would player B wish to engage in the treaty for longest, with highest probability? How much can x and z vary to remain being within 0.05 of this highest probability?
- 3) Identify which pure strategies are redundant for all values of x and z (i.e., never participate in an equilibrium, and thus never have to be considered).
- 4) Which mixed strategies maximise the utility to player A?

- 5) What is the most sensitive ratio of strategies to be affected by changes in x and z ? Use a metric of $abs(a) + abs(b) + abs(c)$ where a, b, c are coefficients in the regression of x, z and $x \times z$.

A reader will note that we treat the under-specifications x and z in different capacities for the five questions. Questions 1, 4 and 5 imply that a modeller has some level of control over the values of x and z , and seeks answers for specific values. Question 2, in a similar manner seeks an exact x and z , but these are the objects of optimisation in order to achieve a result. Question 3 on the other hand seeks an answer irrespective of the under-specification (indicating, potentially, that a modeller has no control over the parameters).

5 Results of our approach for this case study

For our case study and sample queries of Section 4, we employ our approach using the constraints $0.5 \leq x \leq 0.9$ and $0.5 \leq z \leq 0.9$. We have chosen these global constraints after testing and seeing that such a range is achievable by the model, and that we should expect some meaningful, non-trivial results (they are not meant to be indicative of any real scenario). There also happen to be up to 10 mixed Nash equilibrium points in total for each concrete (x, z) pair, over the ranges of acceptable x and z . For sake of argument, we fix the ‘treaty length’ time unit (player B’s strategies depend on) as one year.

Question 1 We ask Z3 to seek the maximum use of 1 inspection only (strategy 1 for player A) across all equilibrium points, and allow x and z to vary independently within their bounds. Ten such equilibria exist.

In the analysis of these five questions, `FunIjx` (for some letter, I and number j) denotes the regression function in SMT for each strategy that contributes towards a Nash equilibrium. The letter I represents the player (either A, the inspector or B, inspected nation) and the number j denotes the strategy of that player. A witness for a satisfiable model is evaluated for the variables of interest at the x and z values given below. We have edited out variables of non-importance (pure strategies used none of the time in this equilibrium, for instance) to save space. We find that 10 such equilibria occur if $z = 0.75$ and $x = 0.9$, maximising the inspector’s strategy 1 (the use of 1 inspection only - `FunA1x`) in each instance to around 48% of the time:

```
[x = 9/10,
z = 3/4,
FunA7x = 4273636400000007368771/81920000000000000000,
FunA1x = 391836359999992690167/81920000000000000000,
FunB7x = 1583308799999994471/81920000000000000000,
FunB1x = 3304345600000002408801/40960000000000000000]
```

Due to space constraints we (arbitrarily) focus on the first equilibrium, for this and all subsequent questions (if applicable).

In the 10 equilibria, player A only ever forms a mixed strategy comprising of pure strategies for 1 inspection and either 7 (as seen here - `FunA7x`) or 8 inspections. The remainder of the variation between the different Nash equilibria that are of maximum use of 1 inspection are due to variety in the treaty length (i.e. change mixed strategies for player B). B can vary mixed strategies between either 2,3,4,5 and 6 year treaty lengths or the full 8 year pure strategy to form the equilibrium.

This sort of analysis could be useful to a decision maker if, for instance, A and B have some control over x and z , and are able to determine their values (perhaps they model ‘costs’ that are incurred, or requirements or commitments to other arms negotiations they can use to trade off against intrusive/expensive inspections). Knowing how these change the structure of the equilibrium formed would give the nations something to negotiate towards in any potential treaty.

Question 2 We maximise the use of running the treaty for 8 years (player B’s strategy 7), over uncertainty x and z , and find that this could occur with probability 0.34 (i.e. 34% of the time), if $x = z = 0.5$. This time, we achieved this in 9 equilibrium points; the first of which is shown and discussed for sake of illustration:

```
[FunB1x = 1312500000000003191/2000000000000000000,
FunA1x = 135420000000005389/4000000000000000000,
FunA7x = 26457999999994571/4000000000000000000,
FunB7x = 6874999999999521/2000000000000000000,
z = 1/2,
x = 1/2]
```

If the nations don’t wish for the treaty to run for 8 years (FunB7x - recall from Figure 3 that their possible strategies start at time unit 2, not 1, hence the 7), their other optimum strategy in this equilibrium would be to run the treaty for either of 2,3,4,5 or 6 years (dependent on the equilibrium and number of inspections, which are determined by the different strategies of A). In the first equilibrium, they would agree on a treaty for 2 years (FunB1x) 66% of the time.

To determine the possible values of x and z such that we are still within 0.05 probability of maintaining this maximum treaty run time of 0.34 (i.e., checking all x and z such that B can go down to using the strategy only 29% of the time), we constrain the solver to look for models with FunB7x = 0.29, and find that this returns a model with $x = z = 0.585$. Running a satisfiability check to confirm that FunB7x’s value does not fall to less than 0.29 within the range $0.5 \leq x \leq 0.585$ or $0.5 \leq z \leq 0.585$, we confirm that this range is acceptable for x and z .

Knowing a range of ‘acceptable’ values of x and z could help negotiators with working towards a treaty between the two nations that is acceptable to both for carrying out the roles of A and B. Although not modelled here, x and z could refer to other ‘sweeteners’ in the treaty that could be used to build confidence in each other in the absence of an available inspection.

Question 3 Here we allow x and z to vary and aim to find a probability greater than 0 that a pure strategy will be included in the mixed strategy, for all equilibria and for all pure strategies. If the SMT solver tells us that this is impossible then the pure strategy’s use ratio has a maximum of 0, and it is never used, for any x and z , and so is thus redundant.

Equilibrium points:		1	2	3	4	5	6	7	8	9	10	Redundant?
Strategies for player A	Holding ___ inspections											
1	1	0.4783502	0.4783502	0.4783502	0.4783502	0.4783502	0.4783502	0.4783502	0.4783502	0.4783502	0.4783502	
2	2	0	0	0	0	0	0	0	0	0	0	0 REDUNDANT
3	3	0	0	0	0	0	0	0	0	0	0	0 REDUNDANT
4	4	0	0	0	0	0	0	0	0	0	0	0 REDUNDANT
5	5	0	0	0	0	0	0	0	0	0	0	0 REDUNDANT
6	6	0	0	0	0	0	0	0	0	0	0	0 REDUNDANT
7	7	0.66145032	0	0.66145032	0	0.66145032	0	0.66145032	0	0.66145032	0	0
8	8	0	0.66145032	0	0.66145032	0	0.66145032	0	0.66145032	0	0.66145032	0
Strategies for player B	___ Treaty units length											
1	2	0.89700995	0.89700995	0	0	0	0	0	0	0	0	0
2	3	0	0	0.89700995	0.89700995	0	0	0	0	0	0	0
3	4	0	0	0	0	0.89700995	0.89700995	0	0	0	0	0
4	5	0	0	0	0	0	0	0.89700995	0.89700995	0	0	0
5	6	0	0	0	0	0	0	0	0	0.89700995	0.89700995	0
6	7	0	0	0	0	0	0	0	0	0	0	0 REDUNDANT
7	8	0.34375033	0.34375033	0.34375033	0.34375033	0.34375033	0.34375033	0.34375033	0.34375033	0.34375033	0.34375033	0

Table 1. Maximum usages of strategies over all $0.5 \leq x \leq 0.9$, $0.5 \leq z \leq 0.9$ for all equilibrium points, highlighting the pure strategies that are redundant - those whose row entries 1-10 are all zeros.

Player A's 1st, 7th and 8th pure strategies are the only ones that lead to equilibrium points within the specified ranges of parameters x and z (so pure strategies 2-6 are redundant). Player B's 6th pure strategy (a treaty that runs for 7 years) is redundant as it never leads to an equilibrium point.

The redundant strategies observed seem to reflect that, given these parameters, the advice to decision makers should be either to push for very many inspections (leading to high assurance, but at high cost) or very few (more trusting, but with lower expense). Our new methodology shows that both approaches to arms control regimes outperform the alternative of seeking a middle ground.

Question 4 We have written a function, `calcpayoffs()`, that calculates payoffs from mixed strategies. It takes the results of a Z3 model (which, themselves are the product of regressions from the *Gambit* models) and calculates the game's expected payoff with those strategies; essentially calculating the perceived utility of the extensive game over the equilibria strategies of the normal game.

We compute that the max utility for player A to arise from any of the equilibria is -0.513281683224. This can occur in a few equilibria, but one example is the following mix of strategies, where x and z were again allowed to vary:

```
[x = 3/4,
 z = 3/4,
 FunA7x = 9185199999999996787/16000000000000000000,
 FunA1x = 1362960000000000639/32000000000000000000,
 FunB7x = 30923999999999991/160000000000000000,
 FunB1x = 6453800000000003921/8000000000000000000]
```

Maximising the utilities over all equilibria and over all possible x and z helps a decision maker choose which equilibrium to strive for in a treaty negotiation, but also gives insight into how x and z affect the utility of the inspection regime. If it is known that they have a relatively higher actual 'cost' (negative pay-off) than $x = z = 0.75$ - where our maximum utility is obtained - it may be useful to know that under these circumstances the maximum utility of the regime won't be attained.

The nations - with the ability to model the utility to A and B explicitly - will be able to 'choose' which is the most important role for them (inspecting or being inspected) and vary strategies accordingly to reflect. Such a decision may be based on whether they trust the other nation going in to the treaty.

Question 5 We allow x and z to vary, whilst maximising each **Sensitivity** function for all pure strategies of A and B.

The most sensitive this can be is 0.6019, using a metric of $abs(a) + abs(b) + abs(c)$ where a, b, c are the regression coefficients of 'regression features' x, z and $x \times z$. These features (plus an intercept) are all those that are possible in a linear regression model of x and z , and are all deemed important enough to be included in the model by testing the effect of removing them from the model one-by-one. This most sensitive value occurs in many equilibria, though x and z both equal 0.5 for all of them. Many other equilibria had much smaller sensitivities.

Sensitivity to variations in x and z highlights the most risky of strategies to a decision maker. Small changes to the true pay-off under these circumstances will have the greatest effect on the choice of strategies, and might change the position of an equilibrium. Knowing such maximal sensitivities is useful to a decision maker as it highlights circumstances where the use of certain pure strategies may be problematic.

6 Future Work

We could explore alternative methods of implementing our linear regression approach, or, indeed replace it completely with an alternative method of capturing the symbolic nature of the under-specified solutions. For instance, bisecting the problem may not be the optimal way of achieving a termination condition in the acceptable regression fitting computation. It may turn out that 'splitting'

the regressions over the ‘boundary points’ is less computationally intensive. Such potential needs exploring.

The study of solution concepts of games that don’t rely on Nash equilibria may be advantageous in this domain. The concept of ‘secure equilibria’ [9], ‘Pandora’s rule’ [22] or any other solution concept may change our ‘solution’ standard. For example, secure equilibria are such that neither player has any incentive to choose a different equilibrium and could add trust between two parties in this domain.

Recent work on Bayesian Belief Networks [5,6] and dynamical systems [7] cover related modelling problems, and said work has focussed on engineering scalable solutions to these complex computations. These approaches could be integrated. In particular, it is clear that a BBN or game theoretic model could sit comfortably within an otherwise deterministic dynamical system, with the possibility of results and workings from the dynamical system driving modelling aspects of the probabilistic models embedded within it. Overall, this could lead to models that encompass more aspects and nuances of real-world arms control regimes.

We are also interested in exploring alternative approaches to uncertainty in game theoretic models; particularly epistemic game theory [16], where there appears to be great overlap in aims as pursued in this paper - although wholly different approaches - to resolving unknowns in models.

7 Conclusions

In this paper, we have explored the analysis of inspection regime models in a game theoretic setting to determine its potential uses in supporting decision making.

We have presented a methodology for dealing with uncertainty in pay-offs assigned to players in non-cooperative, normal form, game theoretic models. This approach generates symbolic results for expected pay-off’s and mixed strategy profiles required to generate Nash equilibria that can be reasoned by a SMT constraint solver. The result is an automated approach that could allow a decision maker to ask pertinent questions of a game theoretic model under uncertainty with some of the data in the model.

Our core contribution has been a methodology to compute with symbolic pay-offs through forming regressions of mixed strategies used to form Nash equilibria. We have established the key concepts behind this methodology and described how it allows us to generalise results: using linear regression of numerical values captured at points over an interval of interest. Encapsulating these results in SMT, which solves these arithmetic and symbolic computations (as well as any other logical constraints the model may have), allows us to check the satisfiability of scenarios overall.

Overall, we have shown how we can generalise models of historic importance in the study of nuclear arms dynamics, whilst effectively addressing the modelling concerns of selecting ‘correct’ pay-off values, and automatically analysing questions of interest for a range of possible values. This allows us to support decision making in particular in our scenario of interest.

We can now provide answers to questions such as *Given uncertainty in what would happen when an inspecting nation wished to hold more inspections than they were permitted to by the host, how long can a treaty be beneficial for both parties in such circumstances?*, and can identify which pure strategies are redundant *for all* cases even under uncertainty. We can also assess which pure strategies are the most sensitive to changes in such under-specifications, providing support to a decision maker on potential strategies to avoid. We believe that analysis of this type was not previously supported; and the capability to do this is useful for scenarios such as the one discussed here as a case study; in supporting decision making, but also more broadly - in gambling, choice games or strategy optimising.

Acknowledgements: The authors from Imperial College London would like to thank AWE for sponsoring a PhD studentship under which the research reported in this paper was carried out.

Open Access of Research Data and Code: The Python and SMT code for the queries and models of this paper and raw SMT analysis results are reported in the public data repository <https://bitbucket.org/pjbeaumont/inmm2016/>

References

1. Avenhaus, R., Von Stengel, B., Zamir, S.: Inspection Games - Chapter 51, Handbook of Game Theory, Vol. 3, North-Holland, Amsterdam, p22, 2002
2. Avenhaus, R., Canty, M. D., Kilgour, D. M., von Stengel, B. and Zamir, S.: Invited review: Inspection games in arms control. European Journal of Operational Research 90, p383-394, 1996
3. Barrett, C., de Moura, L.M., Ranise, S., Stump, A., Tinelli, C.: The SMT-LIB initiative and the rise of SMT - (HVC 2010 award talk). In: Hardware and Software: Verification and Testing - 6th International Haifa Verification Conference, HVC2010, Haifa, Israel, Revised Selected Papers. p. 3, 2010
4. Barrett, C., Tinelli, C.: CVC3. In: Damm, W., Hermanns, H. (eds.) Proceedings of the 19th International Conference on Computer Aided Verification (CAV '07). Lecture Notes in Computer Science, vol. 4590, pp. 298–302. Springer-Verlag (Jul 2007), Berlin, Germany
5. Beaumont, P., Evans, N., Huth, M., Plant, T.: Confidence analysis for nuclear arms control: SMT abstractions of Bayesian Belief Networks, Computer Security – ESORICS 2015, Lecture Notes in Computer Science, Springer, 2015
6. Beaumont, P., Day, E., Evans, N., Haworth, S., Huth, M., Plant, T., Roberts, C.: An in-depth case study: modelling an information barrier in Bayesian Belief Networks, INMM57, 24-28 July, Atlanta, USA, 2016
7. Beaumont, P., Evans, N., Huth, M., Plant, T.: Bounded Analysis of Constrained Dynamical Systems: A Case Study in Nuclear Arms Control, INMM57, 24-28 July, Atlanta, USA, 2016
8. Castillo, L., Dorao, C. A.: Consensual decision-making model based on game theory for LNG processes, Energy Conversion and Management, Volume 64, 2012
9. Chatterjee, K.: Games with secure equilibria, Logic in Computer Science, 2004
10. Dresher, M., A sampling inspection problem in arms control agreements: a game theoretic analysis. Memorandum No. RM?2972?ARPA, The RAND Corporation, Santa Monica, California (1962)
11. McKelvey, R. D., McLennan, A. M., Turocy, T. L. (2014). Gambit: Software Tools for Game Theory, <http://www.gambit-project.org>.
12. de Moura, L.M., Bjørner, N.: Z3: An efficient SMT solver. In: TACAS. pp. 337–340, 2008
13. Myerson, R. B: Game Theory: Analysis of Conflict - Harvard University Press, p. 1. Chapter-preview links, pp. vii-xi, 1991
14. https://nnsa.energy.gov/sites/default/files/Joint_USUK_Report_FINAL.PDF
15. <http://www.nti.org/about/projects/international-partnership-nuclear-disarmament-verification/>
16. Pacuit, Eric and Roy, Olivier, "Epistemic Foundations of Game Theory", The Stanford Encyclopedia of Philosophy (Spring 2016 Edition), Edward N. Zalta (ed.)
17. SageMath, the Sage Mathematics Software System, 2016, <http://www.sagemath.org>.
18. Schelling, T., The Strategy of Conflict, Harvard University Press, reprinted May 1981
19. UK MoD: Ministry of Defence of the United Kingdom. the UK/Norway initiative: report on the UKNI nuclear weapons states workshop, March, 2010
20. Saaty, T. L.: Mathematical models of arms control and disarmament: application of mathematical structures in politics, Publications in operations research, Wiley, 1968
21. von Stengel, B. : Recursive Inspection Games, CoRR, 2014
22. Weizman, M. L.: Optimal search for the best alternative, Econometrica, Vol. 47, No. 3, 1979