Changing the Game: Some Intuitions Derived from Modelling the Security Dilemma as an Infinitely Iterated Game

This paper examines the implications of modelling the security dilemma as an infinitely iterated game. It starts by examining the roots the security dilemma, where actors' efforts to increase their security cause others to lose their security. This leads to a feedback loop where actors accumulate power but do not increase security. Analysing a simple one period model of the security dilemma indicates that actors will necessarily be trapped in the security dilemma. However, expanding the model to better reflect reality yields useful intuitions for the security dilemma, including the key intuitions that the security dilemma can be overcome, even if not permanently, and is likely to be more rapidly overcome if actors interact more frequently.

The Roots of the Security Dilemma

2 Transparency is a choice. Actors choose how much information to disclose about their actions and intentions on a continuum between openness and secrecy[1], where "secrecy means deliberately hiding your actions" (Florini, 1998, p. 50) and openness means providing full disclosure about actions, including the intentions behind the actions[2]. An actor may provide different actors with different levels of transparency for different actions.

3 Knowing an actor's intentions is more important than knowing its actions, because the actor's intentions will guide its future actions. An actor with an aggressive intent poses greater concern than an actor with a non-aggressive intent, even if both take the same action. For example, the addition of one nuclear-tipped ICBM to North Korea's arsenal is more concerning to the United States than the addition of a similar ICBM to the United Kingdom's, because the United States believes that the United Kingdom is non-aggressive.

4 Unfortunately, it is difficult for other actors to determine an actor's intentions. Under external scrutiny, an actor may be forced to reveal its actions. For example, the full extent of

[1] Adapted from (Florini, 1998). *Contra* Florini, this paper eschews transparency/secrecy as the ends of the continuum to avoid any suggestion that ideal transparency is necessarily good.
[2] Where it is able to fully hide its actions, an actor's intentions are automatically also hidden.

Iraq's nuclear weapons program was revealed by the Gulf War, where it had previously been hidden (Florini, 1998, p. 53). However, it cannot be forced to reveal its intentions, because it is harder to verify intentions than actions. For example, prior to agreeing to the Joint Comprehensive Plan of Action ("JCPOA"), Iran had maintained a stockpile of over 25,000 pounds of 20 percent purity enriched uranium (Sanger & Kramer, 2015). It was hard for other actors to verify Iran's declared intention that the "fuel was for a specialty reactor to make medical isotopes" (Sanger & Kramer, 2015) and even if Iran's intention matched its declarations, the United States clearly did not believe Iran as the transfer of the stockpile was a key component of the JCPOA. As such, actors must make their decisions based on their perception of other actors' intentions.

5 Even if an actor is transparent about its intentions, it may not be believed or its behaviour may be otherwise incorrectly interpreted. For example, the People's Liberation Army's cornerstone strategic theory of active defence is rooted in the principle of "striking only after the enemy has struck" (Johnson, 2017, p. 274). However, it is unclear what actions would constitute the first strike; "the first shot on the plane of politics and strategy must be differentiated from the first shot on the plane of tactics" (Johnson, 2017, p. 275). As such, US defence analysts face challenges in how China may interpret active defence and in the worst case, may interpret China as "a revisionist rising power harbouring 'malign' (i.e. expansionist or aggressive) intent" (Johnson, 2017, p. 272). This uncertainty over another actor's intentions is the root of the security dilemma: "the means by which a state tries to increase its security decrease the security of others" (Jervis, 1978, p. 169).

A Game Theoretic View of the Security Dilemma

6 This paper will adapt Tang's framework to describe the security dilemma (Tang, 2009, p. 594). The security dilemma occurs when two actors which do not intend to threaten each other fear each other's intentions, and as a result of this fear, one actor seeks to increase its security by accumulating power. Even if the first actor solely intends to use this additional power for its own security, the act of accumulation may be seen as threatening by the other actor, who in turn will seek to accumulate its own offsetting power. This can lead to a feedback loop in which both actors continue to accumulate power without necessarily

increasing their own security. With sufficient iterations of the feedback loop, the security dilemma may potentially lead to armed conflict.

7 The security dilemma can be modelled through a two-player single period simultaneous Stag Hunt game (Jervis, 1978, p. 171). An actor can choose either to maintain status quo or accumulate power (more broadly, they can choose to cooperate or compete). Both actors would be best off if they maintained status quo, as they would not increase tension yet would maintain their relative security. An actor would be second best off if it were to accumulate power while the other maintains status quo, as it would have a greater relative security. Both actors would be third best off if they both accumulate power, leading to the security dilemma. Finally, an actor would be worst off if it were to maintain status quo while the other accumulates power, as it would have lost relative security. This game is illustrated in Figure 1 below, where the payoff is listed in terms of ranked outcome instead of units of utility[3].

		A	
		Maintain Status Quo	Accumulate Power
B	Maintain Status Quo	1 1	2 4
	Accumulate Power	4 2	3 3

Figure 1 - Stag Hunt

8 The game illustrated above falls into the class of games known as coordination games, which have multiple Nash equilibria[4]. There are two pure-strategy Nash equilibria, where both actors either jointly maintain status quo or jointly accumulate power. Assuming that

[3] In other disciplines, games are typically illustrated in terms of units of utility per action, rather than ranked outcome. As such, care needs to be taken in interpreting such illustrations.

[4] See (Nash, 1951). In short, a Nash equilibrium occurs when neither actor can change their actions unilaterally without becoming worse off. Nash equilibria can be pure strategy, where actors will always choose one strategy to achieve it, or mixed-strategy, where actors choose their strategy according to a probability distribution.

both actors are rational, we should expect each game to result in either joint status quo or joint accumulation. However, it is not immediately clear which of these two equilibria will be chosen for any one game.

9 Harsanyi and Selten's concepts of payoff dominance and risk dominance provide a guide to which equilibrium is likely to be chosen. Under payoff dominance, an equilibrium will dominate another equilibrium if it yields every actor a strictly higher payoff than the other does (Harsanyi & Selten, 1988, p. 356). In contrast under risk dominance, if one actor does not know the other actor's intentions (i.e. which strategy the other actor will choose), an equilibrium will dominate another equilibrium if the strategy an actor needs to achieve it is less risky than the other (i.e. the risk dominant strategy is the strategy with the best worst-case outcome) (Harsanyi & Selten, 1988). In the game above, the payoff dominant equilibrium is joint status quo, while the risk dominant equilibrium is joint accumulation.

10 Achieving the payoff dominant equilibrium requires that all actors are collectively rational (Harsanyi & Selten, 1988, p. 356). Collectively rational actors will "cooperate in pursuing their common interests if the conditions permit them to do so" (Harsanyi & Selten, 1988, p. 356), i.e. where they are certain of the other's intentions. In other words, an actor would not deliberately choose joint accumulation if they were able to achieve joint status quo. This is equivalent to Tang's assertion that a genuine security dilemma can only exist between two actors that "merely want security without intending to threaten the other" (Tang, 2009, p. 594). If conditions permit, we should therefore always expect actors in a security dilemma to "select actions that support a unique payoff-dominant equilibrium" (Straub, 1995, p. 342) and end up with joint status quo.

11 However, if conditions do not permit, we should instead expect that actors should always select the risk dominant strategy of joint accumulation. Where one actor is uncertain of the other actor's intentions it is riskier for it to choose maintaining status quo, because the other actor could choose to accumulate power. Being rational, the other actor can recognise that the first actor will be less likely to maintain status quo, which in turn will make accumulating power more attractive. Through an iterative process of deduction, both actors

recognise that where intentions are uncertain it is always safer to choose to accumulate power, thus leading to the security dilemma.

Changing the Game

12 To escape the security dilemma, actors must be collectively rational. This requires that both actors must be able to determine the other's intentions and that neither actor intends to threaten the other. But in reality, actors' intentions are hard to determine and the model therefore suggests that both actors will necessarily be caught in the security dilemma. How, then, can players achieve collective rationality?

13 One approach is to change the structure of the game. This follows Wagner's argument that the Stag Hunt is an inadequate model of the security dilemma and that "developing more accurate models of the general problem of enforcing agreements in a state of anarchy will help us to better understand why international cooperation is more easily achieved in some areas than others" (Wagner, 1983, p. 330). The security dilemma was first modelled as a game in the 1970s and since then, more conceptual tools have been developed to deal with games with multiple Nash equilibria. Concurrently, better empirical tools such as agent-based modelling exist to verify theoretical models. One example is the theory of rational learning, which is the premise that "decision-making agents update their beliefs about what other agents will do, based on the actions that they have observed" (Levy, 2015, p. 1)[5].

14 To better reflect reality, the security dilemma can be modelled as an infinitely iterated[6] two-player Stag hunt game of imperfect information, where neither actor knows the other's preferences. The game described in para 7 exists in a vacuum, where both actors are presented with a single choice between maintaining status quo or accumulating power, which

[5] In essence, prior to the start of a game, actors assign probabilities to the likelihood of each potential action the opponent may take. Every action the opponent takes changes the probability of a future action happening and reveals something about the true distribution of events. Learning takes place when the actor updates their expected probabilities based on the actions the opponent has taken.

An analogy can be drawn from a situation with two coins being tossed. The potential outcomes are (H, H), (H, T), (T, H), (T, T). Prior to the coins being tossed, the probability of each outcome is 0.25. However, depending on the outcome of the first coin toss, the probability of each outcome changes. If the first coin toss is H, outcomes (T, H) and (T, T) have a probability of 0 while the probability of outcomes (H, T) and (H, H) are 0.5.

[6] Infinite iteration is compatible with reality, since no real-world actor knows when the game is going to stop. Additionally, it allows for easier modelling than finite iteration (c.f. (Snidal, 1985)).

ends the game. Accumulating power is both instantaneous and simultaneous. Both actors know the other's preferences (i.e. joint cooperation is payoff dominant), even if they do not know the other's intentions. In reality of course, this is not the case. Real world actors interact over multiple time periods, do not necessarily know other actors' preferences and accumulating power takes time to achieve.

15 In infinitely iterated games, if actors are able to learn, they will be able to predict the other actor's intentions over a finite period. If actors know their strategies and their beliefs about their opponents' intentions are not entirely false, they should over time converge to a true understanding of the opponent's preferences and intentions[7], leading both actors to effectively[8] choose a Nash equilibrium of the repeated game (Kalai & Lehrer, 1993). The further each actor's starting belief about their opponent's interactions is from the truth, the longer it will take to achieve understanding (Kalai & Lehrer, 1993). However, there are limits to learning. While it is possible for actors to maintain the Nash equilibrium in the short run, i.e. over any fixed length period, maintenance cannot be ensured over the long run (Levy, 2015). Additionally, this result assumes the existence of perfect monitoring, where "before making the choice of a period's action, the player is informed of all the previous actions taken" (Kalai & Lehrer, 1993, p. 1020). If perfect monitoring does not hold, it is possible that the actors may never converge to an understanding of each others' intentions.

16 Empirical evidence suggests that continued repetition of a coordination game can lead to cooperation. Using agent-based modelling, Fang et al have found that after sufficient iteration, cooperative behaviour in a Stag Hunt game can emerge from trial and error learning (Fang, Kimbrough, Valluri, Zheng, & Pace, 2002). This is achieved through modelling an algorithmic agent which starts with arbitrary beliefs about actions that are adapted over time, such that "the tendency to choose an action in a given state is strengthened if it leads to positive results, weakened if the results are unfavourable" (Fang, Kimbrough, Valluri, Zheng, & Pace, 2002, p. 452). Conducting 5000 runs of 500 plays against opponents with fixed

[7] This does not mean that the actor knows the other actor's intentions with certainty; "beliefs may only converge to the truth as time goes by without ever coinciding with it" (Kalai & Lehrer, 1993, p. 1025). However, after sufficient time, the probability that the actor's beliefs match with the other actor's intentions will be arbitrarily close to one (Kalai & Lehrer, 1993).
[8] The actors will only be able to reach an ε-Nash equilibrium, where ε is arbitrarily small.

strategies, “cooperation emerges successfully in matches in which the learner is pitted against relatively trusting opponents” (Fang, Kimbrough, Valluri, Zheng, & Pace, 2002, p. 455) with fixed strategies, such that by the end of each run, cooperation is achieved 100% of the time. Additionally, where learners have longer memories, they are better able to adapt, particularly where the opponent plays a varying strategy. Most importantly, results when both agents are learners suggest that the agents are more able to cooperate than not, with mutual cooperation resulting in about 73% of all runs.

Analysing the Modified Model

17 Analysing the modified model allows us to draw some useful intuitions which can be used as a guide for thinking about the security dilemma. The first intuition is that actors can indeed escape the security dilemma. When first modelled, using the Stag Hunt to illustrate the security dilemma “greatly [exaggerated] the problem of achieving jointly preferred outcomes when there is incomplete knowledge of preference[9]” (Wagner, 1983, p. 345). However, with sufficient time and interaction, actors will be able to discern each other’s benign intentions and cooperate over an arbitrarily long (but not infinite) period.

18 Building on this intuition, defence diplomacy may offers a relatively cheap avenue for actors to interact. While no universal definition of defence diplomacy exists, it is commonly accepted as “activities involving a range of peacetime cooperation between militaries” (Bitzinger, 2011, p. 104). This includes diverse activities ranging from large scale efforts such as conducting joint exercises or multilateral contacts between senior military defence officials to individual scale efforts such as exchanges of personnel or conducting training and capacity building (Bitzinger, 2011, pp. 104-105). Such efforts, while potentially expensive on an absolute scale, may be relatively cheap compared to the cost of arms acquisition. For example, in 2016 the US Department of Defence budgeted for an additional US$1 billion over 5 years for increased training (Department of Defense, 2016), some portion of which would support international participation in Exercise Red Flag. In comparison, the estimated total program cost for the F-35 Joint Strike Fighter as at December 2015 was US$379 billion (Department of Defense, 2015, p. 25).

[9] (Jervis, 1978) and (Wagner, 1983) predate (Harsanyi & Selten, 1988), which makes greater strides towards resolving the security dilemma.

19 Another intuition is that the more opportunities that actors have to play the game, the faster an actor will be able to predict the other actor's intentions[10]. This implies that non-aggressive actors should seek to interact as often as possible, so as to build a shared history from which they can learn. Neither actor knows what the other actor's preferences are. If both actors are truly non-aggressive, their actions will over time disclose their intent and both will be able to cooperate. On the other hand, if one actor is aggressive and the other is not, it is essential for the non-aggressive actor to find this out as soon as possible so that it can accumulate power for its own security[11]. Conversely, the aggressive actor should seek to minimise interaction, so that its aggression is not discovered.

20 Engaging in more frequent defence diplomacy could therefore assist in addressing the security dilemma. Through peaceful interaction, actors can better appraise other actors' intentions, thereby reducing tension and the chance for miscalculation. For example, the ongoing defence diplomacy efforts in ASEAN are likely to have reduced the possibility of intra-ASEAN conflict. Additionally, through building personal relationships between militaries, defence diplomacy can increase comfort between states that the other actor is not aggressive. However, defence diplomacy is not a panacea. It does not guarantee that actors will become friendly – instead, it is a means to allow both actors to rationally assess each other's intentions.

21 A third intuition is that actors can still be caught in the security dilemma for a long time. Even though frequent repeated interactions may address the security dilemma, this process of learning takes time and depends on how sceptical an actor is about the other actor's intentions (i.e. how likely it is that the other actor will compete). As a result, until an actor discerns the other actor's intention, both will remain caught in the security dilemma. One possible way to address this in the real world may be for a highly sceptical actor to consider placing a higher expected probability on the other actor's willingness to cooperate.

[10] (Snidal, 1985) provides a related insight, suggesting that "the longer an issue persists, the greater the impact that future play will have on current choices" (Snidal, 1985, p. 51). A way to exploit this might be to "divide negotiations into larger numbers of smaller sequential steps" (Snidal, 1985, p. 51). However, Snidal's approach is predicated on a finitely iterated game and does not describe the process by which the iteration creates incentives to cooperate.

[11] Even if this leads to a spiral of power accumulation, this is not a security dilemma because the accumulation of power is warranted.

Of course, given that the actor is highly sceptical, it is difficult for the actor to do so. For example, even if Iran is actually non-aggressive to Saudi Arabia, it is hard to imagine Saudi Arabia unilaterally deciding to increase its expected probability that Iran is willing to cooperate and maintain the status quo.

22 The fourth intuition is that present situations are not permanent and that even if actors escape the security dilemma for a while, they may still be caught in future. Both actors may be able to predict over time the other's intentions to within an arbitrary ε and thus be able escape the security dilemma for an arbitrarily long but not infinite period. However, $\varepsilon > 0$ and therefore there still exists the possibility that at some point in the infinite iterations in the game, the actors will incorrectly predict the others' intentions and re-enter the security dilemma. Separately, if the process of learning is still under way, both actors may happen by trial and error to cooperate for extended periods, but revert to competition as part of the learning process. Conversely, of course, the actors may be caught in the security dilemma for a long time, but still be able to escape.

23 Further, the fifth intuition is that where both actors have long memories, they should be able to escape the security dilemma faster than where both actors have shorter memories. Given that perfect monitoring is not achievable in the real world, the learning process may not converge to an understanding of the other actor's intentions and preferences. An actor with a short memory will have a shorter shared history and hence will be less able to learn its opponent's preferences and intentions, and vice versa. This implies that actors with shorter memories should seek to interact even more frequently with other actors, so that they have a greater chance of learning.

24 Finally and somewhat depressingly, the last intuition is that where one actor has a short memory and the other has a long one, the likelihood of escaping the security dilemma depends solely on the actor with a short memory. If the actor with a short memory cannot learn the other's intentions, the actor with a long memory must resort to the risk-dominant strategy and be mired in the security dilemma. If one believes that North Korea, which has had a far longer continuity of government than the United States, is a non-aggressive actor

and if memory is proportional to continuity of government, neither country may never be able to escape the security dilemma.

25 Analysing the model provides useful intuitions for thinking about the security dilemma. However, it should be remembered that these intuitions do not necessarily match reality, because the model does not and cannot capture all real world dynamics. For example, the real world is multilateral, not bilateral, and if aggressive states exist, the legitimate accumulation of power by one state may indeed trigger an unwanted security dilemma. Further work may need to be done to investigate the applicability of the security dilemma in a multilateral world.

Bibliography

Bitzinger, R. A. (2011). U.S. Defence Diplomacy Towards Southeast Asia. In B. Singh, & S. Tan (Eds.), *From "Boots" to "Brogues": The Rise of Defence DIplomacy in Southeast Asia* (pp. 104-115). Singapore: S. Rajaratnam School of International Studies.

Department of Defense. (2015). *Selected Acquisition Report: F-35 Joint Strike Fighter Aircraft.* Washington: US Department of Defence.

Department of Defense. (2016, 2 4). *Remarks by Secretary Carter at Nellis Air Force Base, Nevada*. Retrieved 9 24, 2017, from US Department of Defense: https://www.defense.gov/News/Transcripts/Transcript-View/Article/650577/remarks-by-secretary-carter-at-nellis-air-force-base-nevada/

Fang, C., Kimbrough, S. O., Valluri, A., Zheng, Z., & Pace, S. (2002). On Adaptive Emergence of Trust Behavior in the Game of Stag Hunt. *Group Decision and Negotiation, 11*, 449-467.

Florini, A. (1998). The End of Secrecy. *Foreign Policy, 111*, 50-63.

Glaser, C. L. (1997). The Security Dilemma Revisited. *World Politics, 50*(1), 171-201.

Harsanyi, J. C., & Selten, R. (1988). *A General Theory of Equilibrium Selection in Games.* Cambridge, Mass: MIT Press.

Jervis, R. (1978). Cooperation Under the Security Dilemma. *World Politics, 30*(2), 167-214.

Johnson, J. (2017). Washington's Perceptions and Misperceptions of Beijing's Anti-access Area-Denial (A2-AD) "Strategy": Implications for Military Escalation Control and Strategic Stability. *The Pacific Review, 30*(3), 271-288.

Kalai, E., & Lehrer, E. (1993). Rational Learning Leads to Nash Equilibrium. *Econometrica, 61*(5), 1019-1045.

Lepingwell, J. W. (1987). The Laws of Combat? Lanchester Reexamined. *International Security, 12*(1), 89-134.

Levy, Y. J. (2015). Limits to Rational Learning. *Journal of Economic Theory, 160*, 1-23.

Nash, J. (1951). Non-Cooperative Games. *Annals of Mathematics, 54*(2), 286-295.

Sanger, D. E., & Kramer, A. E. (2015, 12 29). Iran Hands Over Stockpile of Enriched Uranium to Russia. *New York Times*, p. 4.

Snidal, D. (1985). The Game Theory of International Politics. *World Politics, 38*(1), 25-57.

Straub, P. G. (1995). Risk Dominance and Coordination Failures in Static Games. *The Quarterly Review of Economics and Finance, 35*(4), 339-363.

Tang, S. (2009). The Security Dilemma: A Conceptual Analysis. *Security Studies, 18*(3), 587-623.

Wagner, R. (1983). The Theory of Games and the Problem of International Cooperation. *The American Political Science Review, 77*(2), 330-346.