A Game Theoretic Approach to Nuclear Negotiations*

Jonathan Mostovoy
The Hebrew University of Jerusalem and The University of Toronto

September 4, 2015

Abstract

I present a game theoretic modelling of the likely future interactions between Iranian, American, Israeli and European political influencers regarding the recently announced nuclear deal (JCPOA). The model is an extensive form game with complete information amongst countries. I analyse the set of most likely outcomes and their probabilistic subgame equilibria.

1. Introduction

When Ayatollah Khomeini assumed near absolute power over Iran during the Islamic-Iranian Revolution of 1979, immediate conflict arose between this new regime and the United States. Diplomatic relations between the two countries continued to decline, and when the US discovered Iran was developing a nuclear program in 2003 relations worsened even further. Since the Bush Administration, pressure in the form of militaristic threats and sanctions by the US and its European allies for Iran to roll back its nuclear aspirations have been in place. The Obama Administration has showcased an increased effort to avoid a direct confrontation with Iran, and increased its efforts to begin direct diplomatic negotiations with Iran. In July 2015, after nearly two years of such negotiations, a nuclear deal was signed. Such a deal will transform the entire dynamic situation of the Middle East. The most important aspects of the present deal are: limitations on research and development of a nuclear arsenal, allowable but restricted inspections by the Western powers, removal of economic sanctions on Iran, and an increase in Iran's non-nuclear military might.

Since the deal was announced, the Obama administration has lost the support of Congress, but as of September 2, 2015, has enough support in the Senate to allow a presidential veto on congress' override of the deal, which means from the side of the US, the deal will be upheld. However, opinion polls on the decision have waned, lawmakers are up in arms, and even some in the Obama administration have doubts. Since the situation for negotiations and outcomes is representative of a an extensive form game, and with certain intrinsic factors

*I would like to thank the Hebrew University's Economics Department and its Federmann Centre for The Study of Rationality for providing me with the tools to conduct this research. I would also like to specifically thank Professor Eyal Winter for advising me on this research.

such as an absence of mutual trust and each agent acting in a strategic and selfish manner, it seems logical that a game theoretic approach of this recent nuclear deal will highlight certain characteristics of the deal that may assist in the establishment of a truly objective analysis.

2. The Model

2.1. Notation

Consider a model of the interaction between 3 agents: Iran, US + its allies, and Israel as denoted by: $N = \{IRN, US^{+}, ISR\}$. Generic agents, which could be any of our three agents, are represented by i, j and k. X will denote the set of all possible outcomes. θ_{IRN}, $\theta_{US^{+}}$ and θ_{ISR} describe the true preferences of each respective agent over X. The set of all possible types of agent i is denoted at Θ_i. S_i is the strategy set for agent i, $S = \{S_1 \times S_2 \times ... \times S_n\}$ is the set of strategy profiles for all agents, and $S \vdash \Theta$. Utility of i with preference θ_i from outcome $x_i \in X$ is $u_i(\theta_i, x_i|S)$, where u_i is assumed independent of θ_j and θ_k. Initial deviations from a present equilibrium by agent i is given by d_i where $d_i \in S_i$.

2.2. Mechanism

We will assume that all agents begin at the initial node of adhering to the brokered nuclear deal. We will also assume that the nuclear deal is in a temporary Nash equilibrium. By temporary Nash equilibrium (x^* and S^*), I mean that if all agents' profiles are and continue to be what has been publicly disclosed to reach the deal at hand ($\Theta_{public} = \Theta$), then there is no profitable deviation (cheat on the deal) that can produce a larger utility, given by the following inequality:

$$\forall i, x_i \in S \in \Theta : u_i(x_i^*, x_{-i}^*) \geqslant u_i(x_i, x_{-i}^*). \tag{1}$$

If no agent deviates (cheats) on the deal, then the strategies for all agents is know, and thus with a probability 1, the outcome for each stage of the game, x is known.

However, if any agent has lied about their true preferences during negotiations, for example Iran actually prefers a nuclear bomb to the lifting of economic sanctions before 8 years transpire, then $\Theta_{public} \neq \Theta$ and the nuclear deal will only be in equilibrium until the inevitable deviation transpires. When agent i deviates from the deal, a new subgame will become defined and after j and k alter their strategies to account for the $\triangle S_i$, then convergence towards a new equilibrium will eventually occur.

3. The 2 Most Probable Real-Life Subgames

The various probabilities associated to the true (and not publicized) preference types of IRN, US^{+} and ISR have been collected through a process of consulting expert analysis on

the situation.* By averaging the collected opinions, we can determine how probable it was that during negotiations, each agent was being truthful about his intentions (preferences) (in the case of Israel, it will be how probable what they are claiming in the media is to their true preferences). This probability will be denoted as $P(\Theta_{public} = \Theta)$. To highlight the probabilistic responses of each agent to one others' progression of successive likely subgames. $P(Determination)$ will represent each agents' commitment to either acquiring an Iranian nuclear arsenal, or stopping it in the time of an infinite horizon. $P(PowerUse)$ represents the willingness of an agent to use its military power to stop or protect an Iranian nuclear arsenal, & the second figure represents the militaristic strength of each agent, where w is an arbitrarily large positive integer.

List of Probabilities to be Used† ‡			
Country/Alliance	$P(\Theta_{public} = \Theta)$	$P(Determination)$	$P(PowerUse)$
Iran	10%	76%	98% & 2
US + Allies	95%	51%	31% & $w \gg 5$
Israel	95%	96%	85% & 5

In addition to expert opinions, I have also factored in how Iran is a near autocratic regime whose leaders are more concerned about maintaining their political power through strength as opposed to the US's and Israel's leaders who are concerned about maintaining their political powers through democratic representation. This is the reason why $P(Determination)$ for Iran is a direct representation of how much its leaders desire a bomb, and for Israel and the US is partially derived from the most recent public opinion polls regarding how much their citizens want Iran to halt its nuclear aspirations.

3.1. Deal is Upheld Scenario

Now armed with the above probabilities, we can begin to analyse a few subgames. The first calculation would be the most straightforward: $P(DealHT)$, the probability of the deal holding true when no agent deviates.

$$\begin{aligned} q = \min\ & (P_i(\Theta_{public} = \Theta) + (1 - P_i(\Theta_{public} = \Theta))(1 - P_i(Determination))) \\ \therefore P(DealHT) = q &- (1 - (P_j(\Theta_{public} = \Theta))(q)(1 - P_j(Determination)) \\ &- (1 - (P_k(\Theta_{public} = \Theta))(q)(1 - P_k(Determination) \end{aligned} \tag{2}$$

Since we achieve a minimum with Iran enacting as agent i, we can calculate the probability that the deal holds true: $q = 10\% + (90\%)(24\%) = 31.6\%$, and then $P(DealHT) = 31.6\% - (5\%)(31.6\%)(49\%) - (5\%)(31.6\%)(4\%) = 30.76\%$. This shows that there is a probability of 30.76% that the deal is upheld.

*Unfortunately, hard data from which to generate conditional probability estimates is not available. Thus, even through the consultation of experts' opinions, all results through applying the theoretical model outlined in this paper will remain subjective and will likely mirror the biases of the expert. However, even with the subjectiveness of such a model, the use of experts' opinions should reduce the error margin of our calculations.

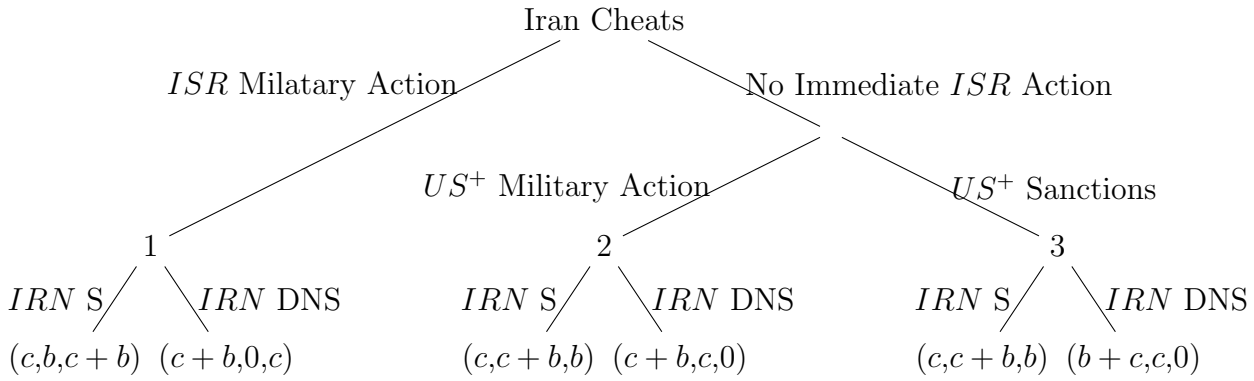
3.2. Iran Cheats Scenario

An actually more probable outcome is that of Iran Cheating $P(IranC)$. This is given:

$$P(IranC) = (P_{IRN}(Determination)(1 - (\Theta_{Public} = \Theta_{IRN})) = (76\%)(90\%) = 68.4\%. \quad (3)$$

However, it must be noted that this probability of cheating is based on any and all types of transgressions on the present deal. For example, if Iran cheats on the margins of the deal by enriching Uranium at 7% as opposed to the allowed 3.67%, it is unlikely that the US^+ and or ISR will launch any attack against Iran. However, if Iran cheats on the deal in a way that will guarantee itself a nuclear arsenal, militaristic threats and economic sanctions will surely be put into action. Furthermore, we will assume that Israel will always have the opportunity to make a decision before the US^+ due to its large $P(Determination)$.

The following diagram is the extensive form game of this situation, where, IRN S = Iranian stops its breach of the deal, and IRN DNS = Iran continues its breach of the deal and game is iterated again for a finite number of plausible iterations; however, payoffs for ISR and US^+ to halt Irans' deal cheating is increased. c = costs incurred reaching a node for i, b = benefit of reaching a particular node for i, Payouts follow: $(x_{IRN}, x_{US^+}, x_{ISR})$. Furthermore, each pair of alternative decisions for i is accompanied with probabilities: $P(Decision1) + P(Decision2) = 1$. $c_{1,i}$ is i's costs from being at node 1.



Since $P_{US^+}(PowerUse) \gg P_{ISR}(PowerUse)$, then $\forall i, c_i,\ c_{2,i} > c_{1,i} > c_{3,i}$, also $\forall i,\ b|(IRNS) = q_i$ and $b|(IRNDNS) = t_i$ where t and q are constants. To find a solution for this situation, we must define the Θ's of each player and deduce from this their S's. Θ will be defined as:

$$\forall i, \Theta_i \in [x_i^* \succeq x_i \iff x_i^* \geq x_i \quad |c \in x_i, b \in x_i] \quad (4)$$

Now, we may define the strategies for player each agent. In the case of Iran, if x_{IRN}^* is $> \sum c_{IRN} + b_{IRN}$ for all future iterations of the game, then Iran will choose $IRN\ S$ in this iteration of the game and the full game will terminate. This then defines the S's of both US^+ and ISR. In the first iteration of the game, Israel will not choose to strike as it can rely on the US^+ to invoke its sanctions as the US^+ much prefers sanctions to military involvement.

If enduring the military attacks by Israel and then the US^+ comes at a greater payoff to Iran than does stopping its own nuclear arsenal development, then Iran will never halt its transgressions; however, the possibility of this being true is negligible. Thus, if the likely reverse is true, then Iran will wait until sanctions are imposed on it again, and then decide to apologize for its transgressions and revert back to the original deal's terms. The other slightly probable outcome is if due to the $P(Determination)$ of the US^+ fails to produce an incentive for it to attack Iran, then Israel will order a military strike and the game will terminate after this action.

In summary, with a probability of 68.4%, Iran will cheat on this nuclear deal. Due to the structure of the present deal, Iran has the option of marginally cheating on the deal and bear little consequences, this is also where the high probability of cheating is derived from. However, if Iran decides to significantly cheat on this deal, sanctions will be set in place, and based off the costs to Iran of the sanctions and two possible military attacks, it will, with a very large probability, decide to stop such cheating after either sanctions are once again imposed on it, and if not after an Israeli military attack.

4. Inisghts From this Analysis

Insights from this game theoretic analysis has highlighted a few interesting characteristics about the Iranian nuclear deal:

1. There is a 30.76% probability that the deal is upheld until the deal's expiration.

2. Should Iran decide to cheat on the deal, it will likely be cheating on the margins of the deal that will not trigger a response from US+ or Israel. However, there is an additional possibility that Iran greatly cheats on the deal triggering sanctions to be brought back. If the US^+ decides not to attack Iran after they refuse to stop their cheating, which is 49% probable given already being at this decision, then Israel will be pushed to attack Iran. Together, one of the scenarios outlined above will happen with a probability of 68.4%.

3. This nuclear deal truly is a gamble in the sense that the range of likely outcomes varies from cooperation to military attacks. Caution is most definitely justified over the implementation of this nuclear deal between Iran and the West.

5. Bibliography

1. Halem, H. (2013, October 13). Statistical Diplomacy: A Bayesian Inference Tool for the P5 1 Iran Negotiations. Retrieved August 20, 2015.

2. IDF more ready than ever to strike Iran, security official says. (2015, August 25). Times of Israel.

3. Iran militia chief: Destroying Israel is nonnegotiable. (2015, March 31). Retrieved August 12, 2015.

4. Osborne, M., & Rubinstein, A. (1994). A Course in Game Theory. Cambridge, MIT.

5. Schelling, T. (1960). The strategy of conflict. Cambridge: Harvard University Press.

6. Winter, E. (2015, March 25). What Obama Could Learn About Negotiating With Iran From My $2,000 Used Car. Time. Retrieved July 17, 2015.