## Appendix 2: The Computation of Perfect Bayesian Equilibria*

A game may be defined in normal, extensive, or network form. In the first case, one will refer to the equivalent extensive form where each side has a single turn represented by an information set and where each strategy is a move available at that information set. The network form allows cycling within the game form. Moves that can be repeated are called "recurrent." A move $x$ is "discounted" if there exists $\mathrm{d}_x$ $(0 < \mathrm{d}_x < 1)$ such that any payoff resulting from completing the move is multiplied by $\mathrm{d}_x$ when viewed as an expected payoff. Such a discount factor is interpreted as a probability of completing the move once it's been chosen. The most general formulation of equilibrium conditions requires two assumptions:

*Assumption 1*: Any cycle must contain at least one discounted move.

The second assumption requires the following:

*Definition 1*: A node S is called a "source" for an information set $\mathcal{I}$ if for any node $\mathrm{N} \in \mathcal{I}$, any cycle containing N also contains S.

*Assumption 2*: Any information set has a source.

A strategy profile is a set of probability distributions over the moves available at each information set. If $x$ is a move available at the information set $\mathcal{I}$ (denoted $x \prec \mathcal{I}$) then $p_x$ is its probability in the strategy profile. There must therefore exist $p_x \geq 0$ such that $\sum_{x \prec \mathcal{I}} p_x = 1$. And a belief system is a set of probability distributions over the nodes of each information set $\mathcal{I}$. If $\mathrm{N} \in \mathcal{I}$ is a node then there must be a belief (to be at N) $\beta_{\mathrm{N}} \geq 0$ such that $\sum_{\mathrm{N} \in \mathcal{I}} \beta_{\mathrm{N}} = 1$. A perfect bayesian equilibrium (PBE) is a pairing of a strategy profile and a belief system such that the strategies are sequentially rational given the beliefs and the beliefs are consistent with the strategies through Bayes Law whenever possible. The goal here is to show that such a PBE is the solution of an equation of the form $\mathrm{F(X)} = 0$ where X is the vector made up of the probabilites $p_x$ and the beliefs $\beta_{\mathrm{N}}$. The components of F express the conditions of sequential rationality and Bayesian consistency. To obtain them one must first construct expected payoffs for each move and hit probabilities for each node. These involve the (partial) transition matrix T defined over the non-final nodes of the game form as well as the "instant payoff" vectors for each player.

We assume an arbitrary order for the non-final nodes of the game. If node $\mathrm{N}_i$ can lead to node $\mathrm{N}_j$ by a move $x$ then $\mathrm{d}_x p_x$ is the probability of that move and it is in the position $(i, j)$ in matrix T. It accounts for the probability $p_x$ that the move is chosen and the probability $\mathrm{d}_x$ that it is completed when it is discounted. For each move $y$ available at node $\mathrm{N}_i$ that yields the payoff $u_y^{\mathrm{A}}$ to player A, he can expect $\mathrm{d}_y p_y u_y^{\mathrm{A}}$ from that move at that node. The instant payoff vector $U^{\mathrm{A}}$ for A has its $i^{th}$ component given by $\sum_{y \prec \mathrm{N}_i} p_y u_y^{\mathrm{A}}$ where the sum is taken over all moves $y$ available at $\mathrm{N}_i$. One can easily prove

*Lemma 1*: Under assumption 1, $(\mathrm{I} - \mathrm{T})$ is non-singular.

The expected utility vector $E_{\mathrm{A}}$ for player A therefore satisfies

$$E^{\mathrm{A}} = \mathrm{T}E^{\mathrm{A}} + U^{\mathrm{A}}$$

or $$E^{\mathrm{A}} = (\mathrm{I} + \mathrm{T} + \mathrm{T}^2 + .. + \mathrm{T}^n + ..)U^{\mathrm{A}} = (\mathrm{I} - \mathrm{T})^{-1}U^{\mathrm{A}} \quad (1)$$

This vector gives the expected payoff for A of reaching each (non-final) node $\mathrm{N}_i$.

Initial beliefs $b$ are typically a trivial probability distribution where a "start node" has probability 1 and all others have probability $0$. As the game is played according to the strategy profile, some nodes are hit at various turns with various probabilities in Markov chain fashion. The entry $\pi_{ij}^n$ in position $(i, j)$ in the matrix $\mathrm{T}^n$ is the probability of moving from node $\mathrm{N}_i$ to node $\mathrm{N}_j$ in exactly $n$ turns. If $\mathrm{N}_i$ is a source for the information set $\mathcal{I}$ that can be reached in the given strategy profile and if $\mathrm{N}_j$ is any node in $\mathcal{I}$ then the ratio $\rho_{ij} = \alpha_j \div \alpha_i$ of the $j^{th}$ to the $i^{th}$ coordinate in the vector

$$\alpha = b(\mathrm{I} + \mathrm{T} + \mathrm{T}^2 + .. + \mathrm{T}^n + ..) = b(\mathrm{I} - \mathrm{T})^{-1}$$

is the total probability of moving directly (without cycling) from $\mathrm{N}_i$ to $\mathrm{N}_j$. Bayesian updating thus yields updated beliefs for $\mathrm{N}_j$

$$\beta_j = \rho_{ij} \div \sum_{\mathrm{N}_k \in \mathcal{I}} \rho_{ik} = \alpha_j \div \sum_{\mathrm{N}_k \in \mathcal{I}} \alpha_k$$

If an information set cannot be reached in the given profile then all $\alpha_k$ are nil and the beliefs can be arbitrary in $\mathcal{I}$. In all cases, beliefs $\beta_j$ will satisfy

$$g_j(\mathrm{X}) = \beta_j \sum_{\mathrm{N}_k \in \mathcal{I}} \alpha_k - \alpha_j = 0 \quad (2)$$

Given beliefs in (2) and expected payoffs in (1), all resulting from the strategy profile, the initial beliefs, and the players' payoffs, we can now define the expected payoff of each move $x$ at any information set $\mathcal{I}$. Since the move is available at each node of $\mathcal{I}$ we must obtain the expected payoff at any such node and weigh them according to beliefs. If at node $\mathrm{N}_i$ the move is final its expected payoff is merely $E_{ix}^{\mathrm{A}} = p_x u_x^{\mathrm{A}}$. If it is non-final and leads to node $\mathrm{N}_j$ then it reads $E_{ix}^{\mathrm{A}} = p_x(u_x^{\mathrm{A}} + \mathrm{d}_x E_j^{\mathrm{A}})$ where $\mathrm{d}_x$ is the discount factor, if any. The expected payoff of move $x$ at $\mathcal{I}$ then reads

$$E_x^{\mathrm{A}} = \sum_{\mathrm{N}_i \in \mathcal{I}} E_{ix}^{\mathrm{A}} \tag{3}$$

Given the strategy profile and initial beliefs, such expected payoffs are well defined for all moves (although they are not unique if the information set is off the equilibrium path since beliefs are then arbitrary). In any case, at each information set $\mathcal{I}$ one can then define a best reply move $x^*$ by $E_{x^*}^{\mathrm{A}} = \max\{E_x^{\mathrm{A}} : x \prec \mathcal{I}\}$. Sequential rationality at $\mathcal{I}$ then requires for all $x$

$$f_x(\mathrm{X}) = p_x(E_{x^*}^{\mathrm{A}} - E_x^{\mathrm{A}}) = 0 \tag{4}$$

Indeed, the move $x$ can only have positive probability if it yields a maximum expected payoff at $\mathcal{I}$ so that either $E_{x^*}^{\mathrm{A}} - E_x^{\mathrm{A}} = 0$ or $p_x = 0$ (or both).

The collection of equations $\{f_x(\mathrm{X}) = 0$ and $g_j(\mathrm{X}) = 0\}$ can now be denoted $\mathrm{F}(\mathrm{X}) = 0$ as announced. The fact that a solution always exists for this equation results from a fixed point argument similar to that obtained in the Nash theorem. We now turn to computation.

The system $\mathrm{F}(\mathrm{X})$ is clearly continuous in the interior of the action space (where all $p_x > 0$ and can be continuously extended to the boundary (thus defining beliefs off the equilibrium path by a limit process). But because the best reply $x^*$ can be discontinuous, it is only differentiable by pieces (within the "regular" regions where the best reply is unique). This may be an impediment in the design of an efficient numerical method. However, it is not difficult to write an equivalent set of smooth equations. For instance, following Nash's original insight, one could write

$$v_x = \max\{0, E_x^{\mathrm{A}} - \sum_{y \prec \mathcal{I}} p_y E_y^{\mathrm{A}}\}$$

where $x$ is a move by A and the sum is taken over all moves available to A at the same information $\mathcal{I}$ as $x$. Then, the transformation

$$p'_x = \frac{p_x + v_x^2}{1 + \sum_{y \prec \mathcal{I}} v_y^2}$$

is similar to that used in Nash's classical existence proof that relies on a fixed point argument (where all $p'_x = p_x$). Now, $f_x$ can be replaced by the differentiable

$$\varphi_x(\mathrm{X}) = p'_x - p_x \tag{5}$$

Regardless of whether one uses (5) or (6) to represent sequential rationality, it remains necessary to ensure that the $p_x$ indeed define true probabilities. One could simply add constraints to that effect. But a simple and effective device is to introduce a new set of real variables $z_x \in \mathbb{R}$ and to let, at each information set $\mathcal{I}$:

$$p_x = \alpha(z_x) \div \sum_{y \prec \mathcal{I}} \alpha(z_y)$$

where $\alpha$ is a 1-1 differentiable function from $\mathbb{R}$ into $(0,\infty)$. For instance, $\alpha(z) = e^z$. The $p_x$ provide strictly positive probability distributions and therefore make Bayesian updating always possible (i.e., the equilibrium is actually sequential). The unknowns $z_x$ are then subsumed into a vector Z and the resulting equations read $\mathrm{G(Z)} = 0$.

One can then define the Jacobian matrix $\mathrm{D_Z G(Z)}$ and consider the differential system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Z}(t) = -[\mathrm{D_Z G(Z)}]^{-1}.\mathrm{G(Z)} \qquad (6)$$

The function

$$\mathcal{L}(\mathrm{Z}) = [\mathrm{G(X)}]^{\mathrm{T}}.\mathrm{G(Z)} \geq 0$$

is continuously differentiable and satisfies along the solution curve of (6)

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}(\mathrm{Z}(t)) = [\mathrm{D_Z G}(\mathrm{Z}(t)).\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Z}(t)]^{\mathrm{T}}.[\mathrm{D_Z G}(\mathrm{Z}(t)).\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Z}(t)]$$
$$= -2[\mathrm{G}(\mathrm{Z}(t))]^{\mathrm{T}}.\mathrm{G}(\mathrm{Z}(t)) = -2\mathcal{L}(\mathrm{Z}(t))$$

or $\quad \mathcal{L}(\mathrm{Z}(t)) = \mathcal{L}(\mathrm{Z}_0)e^{-2t}$

So, $\mathcal{L}$ is a Lyapunov function for the system and it decreases exponentially along the solution curves of (6). $\mathrm{Z}(t)$ therefore converges to a solution of $\mathrm{G(Z)} = 0$ as long as it doesn't get "caught" in regions where the Jacobian matrix would be degenerate (a known limitation of the Newton method).

Computational experience indicates that the numerical solution of (6) usually converges for arbitrary random initial $\mathrm{Z}_0$. Indeed, the use of equation (4) is just as effective as that of (5) despite the discontinuities in $\mathrm{D_X F}$.

In order to also obtain perfect bayesian equilibria that are not sequential one may limit the implementation of the above techniques to some facets of the total action space (where some probabilities are set to zero). Deciding which facet to serach can be the result of a process akin to the elimination of dominated strategies in normal form games. Finally, to ensure well defined beliefs off the equilibrium path only may involve some artificial nodes leading to such information sets. In practice, several ad-hoc devices provide rather exhaustive results.