Plans Are Worthless but Planning Is Everything: A Theoretical Explanation of Eisenhower's Observation

Angel F. Garcia Contreras, Martine Ceberio, and Vladik Kreinovich

Department of Computer Science, University of Texas at El Paso
El Paso, TX 79968, USA
afgarciacontreras@miners.utep.edu,
mceberio@utep.edu, vladik@utep.edu

Abstract. The 1953-1961 US President Dwight D. Eisenhower emphasized that his experience as the Supreme Commander of the Allied Expeditionary Forces in Europe during the Second World War taught him that "plans are worthless, but planning is everything". This sound contradictory: if plans are worthless, why bother with planning at all? In this paper, we show that Eisenhower's observation has a meaning: while directly following the original plan in constantly changing circumstances is often not a good idea, the existence of a pre-computed original plan enables us to produce an almost-optimal strategy – a strategy that would have been computationally difficult to produce on a short notice without the pre-existing plan.

1 Introduction: Eisenhower's Seemingly Paradoxical Observation

Eisenhower's observation. Dwight D. Eisenhower, the Supreme Commander of the Allied Expeditionary Forces in Europe during the Second World War and later the US President, emphasized that his war experience taught him that "plans are worthless, but planning is everything"; see, e.g., [1].

At first glance, this observation seems paradoxical. At first glance, the Eisenhower's observation sounds paradoxical: if plans are worthless, why bother with planning at all?

What we do in this paper. In this paper, we show that this Eisenhower's observation has a meaning. Namely, it means that:

- while following the original plan in constantly changing circumstances is often not a good idea,
- the existence of a pre-computed original plan enables us to produce an almost-optimal strategy (a strategy that would have been computationally difficult to produce on a short notice without the pre-existing plan).

2 Analysis of the Problem

Rational decision making: a brief reminder. According to decision making theory, decisions by a rational decision maker can be described as maximize the value a certain function known as utility; see, e.g., [3, 4]. In financial situations, when a company needs to make a decision, the overall profit can be used as the utility value; in more complex situations, the utility function combines different aspects of gain and loss related to different decisions.

Let us describe this in precise terms. Let x denote a possible action, a describes the situation, and let $u(x, a)$ denote the utility that results from performing action x in situation a.

To describe a possible action, we usually need to describe the values of several different quantities. For example, a decision about a plant involves selecting amount of gadgets of different type manufactured at this plant – and maybe also the parameters characterizing these gadgets. Let us denote the parameters describing an action by $x_1, \ldots, x_n$. In these terms, an action can be characterized by the tuple $x = (x_1, \ldots, x_n)$.

Similarly, in general, we need several different quantities to describe a situation, so we will describe a situation by a tuple $a = (a_1, \ldots, a_m)$.

In these terms, what is planning. Let $\widetilde{a}$ describe the original situation. Based on this situation, we come up with an action $\widetilde{x}$ that maximizes the corresponding utility: $u(\widetilde{x}, \widetilde{a}) = \max\limits_x u(x, \widetilde{a})$. Computing this optimal action $\widetilde{x}$ is what we usually call *planning*.

Situations change. At the moment when we need to start acting, the situation may have changed in comparison with the original situation $\widetilde{a}$, to a somewhat different situation a. Let us denote the corresponding change by $\Delta a \stackrel{\text{def}}{=} a - \widetilde{a}$. In terms of this difference, the new situation takes the form $a = \widetilde{a} + \Delta$.

A not-always-very-good option: applying the original plan to the new situation. One possibility is to simply ignore the change, and apply the original plan $\widetilde{x}$ – which was optimal for the original situation $\widetilde{a}$ – to the new situation $a = \widetilde{a} + \Delta a$.

This plan is, in general, not optimal for the new situation. Thus, in comparison to the actually optimal plan x^{opt} for which

$$u(x^{\text{opt}}, \widetilde{a} + \Delta a) = \max_x u(x, \widetilde{a} + \Delta a),$$

we lose the amount $L_0 \stackrel{\text{def}}{=} u(x^{\text{opt}}, \widetilde{a} + \Delta a) - u(\widetilde{x}, \widetilde{a} + \Delta a)$.

A better option: trying to modify the original plan. Why cannot we just find the optimal solution for the new situation? Because optimization is, in general, an NP-hard problem (see, e.g., [2, 5]), meaning that it is not possible to find the exact optimum in reasonable time.

What we can do is try to use some feasible algorithm – e.g., solving a system of linear equations – to replace the original plan $\widetilde{x}$ with a modified plan $\widetilde{x} + \Delta x$.

Due to NP-hardness, this feasibly modified plan is, in general, not optimal, but we hope that the resulting loss $L_1 \stackrel{\text{def}}{=} u(x^{\text{opt}}, \widetilde{a} + \Delta a) - u(\widetilde{x} + \Delta x, \widetilde{a} + \Delta a)$ is much smaller than the loss L_0 corresponding to the use of the original plan $\widetilde{x}$.

What we do in this paper. In this paper, we analyze the values of both losses and we show that indeed, L_1 is much smaller than L_0. So, in many situations, even if the loss L_0 is so large that the corresponding strategy (of directly using the original plan) is worthless, the modified plan may leads to a reasonably small loss $L_1 \ll L_0$ – thus explaining Eisenhower's observation.

Estimating L_0. We assume that the difference Δa is reasonably small, so the corresponding difference in action $\Delta x^{\text{opt}} \stackrel{\text{def}}{=} x^{\text{opt}} - \widetilde{x}$ is also small. We can therefore expand the expression for the loss L_0 in Taylor series and keep only terms which are linear and quadratic with respect to Δx. Thus, we get

$$L_0 = u(x^{\text{opt}}, \widetilde{a} + \Delta a) - u(x^{\text{opt}} - \Delta x^{\text{opt}}, \widetilde{a} + \Delta a) =$$

$$\sum_{i=1}^{n} \frac{\partial u}{\partial x_i}(x^{\text{opt}}, \widetilde{a} + \Delta a) \cdot \Delta x_i^{\text{opt}} +$$

$$\frac{1}{2} \cdot \sum_{i=1}^{n} \sum_{i'=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_{i'}}(x^{\text{opt}}, \widetilde{a} + \Delta a) \cdot \Delta x_i^{\text{opt}} \cdot \Delta x_{i'}^{\text{opt}} + o((\Delta a)^2).$$

By definition, the action x^{opt} maximizes the utility $u(x, \widetilde{a} + \Delta a)$. Thus, we have $\dfrac{\partial u}{\partial x_i}(x^{\text{opt}}, \widetilde{a} + \Delta a) = 0$, and the above expression for the loss L_0 takes the simplified form

$$L_0 = \frac{1}{2} \cdot \sum_{i=1}^{n} \sum_{i'=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_{i'}}(x^{\text{opt}}, \widetilde{a} + \Delta a) \cdot \Delta x_i^{\text{opt}} \cdot \Delta x_{i'}^{\text{opt}} + o((\Delta a)^2). \qquad (1)$$

The values Δx_i^{opt} can be estimated from the above condition

$$\frac{\partial u}{\partial x_i}(x^{\text{opt}}, \widetilde{a} + \Delta a) = \frac{\partial u}{\partial x_i}(\widetilde{x} + \Delta x^{\text{opt}}, \widetilde{a} + \Delta) = 0.$$

Expanding this expression in Taylor series in terms of Δx_i and Δa_j and taking into account that $\dfrac{\partial u}{\partial x_i}(\widetilde{x}, \widetilde{a}) = 0$ (since for $a = \widetilde{a}$, the utility is maximized by the action $x = \widetilde{x}$), we conclude that for every i, we have

$$\sum_{i'=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_{i'}}(\widetilde{x}, \widetilde{a}) \cdot \Delta x_{i'}^{\text{opt}} + \sum_{j=1}^{m} \frac{\partial^2 u}{\partial x_i \partial a_j}(\widetilde{x}, \widetilde{a}) \cdot \Delta a_j + o(\Delta x, \Delta a) = 0.$$

Thus, the first approximation Δx_i to the values Δx_i^{opt} can be determined as a solution to a system of linear equations:

$$\sum_{i'=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_{i'}}(\widetilde{x}, \widetilde{a}) \cdot \Delta x_j = -\sum_{j=1}^{m} \frac{\partial^2 u}{\partial x_i \partial a_j}(\widetilde{x}, \widetilde{a}) \cdot \Delta a_j. \qquad (2)$$

A solution to a system of linear equations is a linear combination of the right-hand sides. Thus, the values Δx_i are a linear function of Δa_j. Substituting these linear expressions into the formula (1), we conclude that *the loss L_0 is a quadratic function of Δa_j*, i.e., that $L_0 = \sum\limits_{j=1}^m \sum\limits_{j'=1}^m k_{jj'} \cdot \Delta a_j \cdot \Delta a_{j'} + o((\Delta a)^2)$ for some coefficients $k_{jj'}$.

Estimating L_1. In the previous section, we considered what happens if we use the original plan $\widetilde{x}$ – which was optimal in the original situation $\widetilde{a}$ – in the changed situation $a = \widetilde{a} + \Delta a$. Since the original plan is optimal only for the original situation, but not for the new one, using this not-optimal plan leads to the loss L_0, a loss which we estimated as being quadratic in terms of Δa.

To decrease this loss, we need to update the action x. As we have already mentioned, exactly computing the optimal action $x^{\rm opt}$ is, in general, an NP-hard – i.e., computationally intractable – problem. However, as we have also mentioned, the first approximation Δx_i to the desired difference $\Delta x^{\rm opt}$ – and thus, the first approximation to the newly optimal solution $x^{\rm opt}$ – can be obtained by solving a system of linear equations (2).

The system (2) of linear equations is feasible to solve. Thus, it is reasonable to consider using the action $x^{\rm lin} = \widetilde{x} + \Delta x$ instead of the original action $\widetilde{x}$. Let us estimate how much we lose if we use this new action $x^{\rm lin}$ instead of the optimal action $x_i^{\rm opt}$.

The fact that the difference Δx is the first approximation to the optimal difference $\Delta x^{\rm opt}$ means that we can write $\Delta x^{\rm opt} = \Delta x + \delta x$, where the remaining term $\delta x \stackrel{\rm def}{=} \Delta x^{\rm opt} - \Delta x = x^{\rm opt} - x^{\rm lin}$ is of second order in terms of Δx and Δa: $\delta x = O((\Delta x)^2, (\Delta a)^2)$. Since in the first approximation, Δx has the same order as Δa, we thus get $\delta x = O((\Delta a)^2)$.

The loss L_1 of using $x^{\rm lin} = x^{\rm opt} - \delta x$ instead of $x^{\rm opt}$ is equal to the difference

$$L_1 = u(x^{\rm opt}, \widetilde{a} + \Delta a) - u(x^{\rm lin}, \widetilde{a} + \Delta a) = u(x^{\rm opt}, \widetilde{a} + \Delta a) - u(x^{\rm opt} - \delta x, \widetilde{a} + \Delta a).$$

If we expand this expression in δx and keep only linear and quadratic terms, we conclude that

$$L_1 = \sum_{i=1}^n \frac{\partial u}{\partial x_i}(x^{\rm opt}, \widetilde{a} + \Delta a) \cdot \delta x_i +$$

$$\frac{1}{2} \cdot \sum_{i=1}^n \sum_{i'=1}^n \frac{\partial^2 u}{\partial x_i \partial x_{i'}}(x^{\rm opt}, \widetilde{a} + \Delta a) \cdot \delta x_i \cdot \delta x_{i'} + o((\delta x)^2).$$

Since $x^{\rm opt}$ is the action that, for $a = \widetilde{a} + \Delta a$, maximizes utility, we get

$$\frac{\partial u}{\partial x_i}(x^{\rm opt}, \widetilde{a} + \Delta a) = 0.$$

Thus, the expression for L_1 gets a simplified form

$$L_1 = \frac{1}{2} \cdot \sum_{i-1}^n \sum_{i'=1}^n \frac{\partial^2 u}{\partial x_i \partial x_{i'}}(x^{\rm opt}, \widetilde{a} + \Delta a) \cdot \delta x_i \cdot \delta x_{i'} + o((\delta x)^2).$$

We know that the values δx_i are quadratic in Δa; thus, we conclude that for the modified action, *the loss L_1 is a 4-th order function of Δa_j*, i.e., that

$$L_1 = \sum_{j=1}^{m} \sum_{j'=1}^{m} \sum_{j''=1}^{m} \sum_{j'''=1}^{m} k_{jj'j''j'''} \cdot \Delta a_j \cdot \Delta a_{j'} \cdot \Delta a_{j''} \cdot \Delta a_{j'''} + o((\Delta a)^5)$$

for some coefficients $k_{jj'j''j'''}$.

3 Conclusions

We conclude that:

- the loss L_0 related to using the original plan is quadratic in Δa, while
- the loss L_1 related to using a feasibly modified plan is of 4th order in terms of Δa.

For reasonably small Δa, we have $L_1 \sim (\Delta a)^4 \ll L_0 \sim (\Delta a)^2$.

Let $\varepsilon > 0$ be the maximum loss that we tolerate. Since $L_1 \ll L_0$, we have three possible cases: (1) $\varepsilon < L_1$, (2) $L_1 \le \varepsilon \le L_0$, and (3) $L_0 < \varepsilon$. In the first case, even using the modified action does not help. In the third case, the change in the situation is so small that it is Ok to use the original plan $\widetilde{x}$.

In the second case, we have exactly the Eisenhower situation:

- if we use the original plan $\widetilde{x}$, the resulting loss L_0 much larger than we can tolerate; in this sense, the original plan is worthless;
- on the other hand, if we feasible modify the original plan into x^{lin}, then we get an acceptable action.

So, we indeed get a theoretical justification of Eisenhower's observation.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award "UTEP and Prudential Actuarial Science Academy and Pipeline Initiative" from Prudential Foundation.

References

1. D. Eisenhower, "A speech to the National Defense Executive Reserve Conference in Washington, D.C., November 14, 1957", in: D. Eisenhower, *Public Papers of the Presidents of the United States,* National Archives and Records Service, Government Printing Office, 1957, p. 818.
2. V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1998.
3. R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.
4. H. T. Nguyen, O. Kosheleva, and V. Kreinovich, "Decision making beyond Arrow's 'impossibility theorem', with the analysis of effects of collusion and mutual attraction", *International Journal of Intelligent Systems*, 2009, Vol. 24, No. 1, pp. 27–47.
5. P. Pardalos, *Complexity in Numerical Optimization*, World Scientific, Singapore, 1993.