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
500 W. University
El Paso, Texas 79968, USA
afgarciacontreras@miners.utep.edu,
mceberio@utep.edu@utep.edu, vladik@utep.edu

1. Eisenhower's Observation

- Dwight D. Eisenhower was:
  - the Supreme Commander of the Allied Expeditionary Forces in Europe during WW2
  - and later the US President.
- He emphasized that his war experience taught him that "plans are worthless, but planning is everything".
- At first glance, this sounds paradoxical: if plans are worthless, why bother with planning at all?
- In this paper, we show that this Eisenhower's observation has a meaning:
  - 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.

2. 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.
- E.g., in financial situations, when a company needs to make a decision, the overall profit can be used as utility.
- To describe a possible action x, we usually need to describe the values of several quantities $x_1, \ldots, x_n$.
- E.g., a decision about a plant involves selecting amounts x_i of manufactured gadgets of different type.
- Similarly, we need several quantities $a_1, \ldots, a_m$ to describe a situation.
- Let $u(x, a)$ denote the utility that results from performing action x in situation a.

3. 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:

$$u(\widetilde{x},\widetilde{a})=\max_x u(x,\widetilde{a}).$$

- Computing this optimal action $\widetilde{x}$ is what we usually call *planning*.
- When we need to start acting, the situation may have changed to $a\neq\widetilde{a}$.
- Let us denote the corresponding change by $\Delta a\stackrel{\text{def}}{=}a-\widetilde{a}$, then $a=\widetilde{a}+\Delta a$.

4. Options

- One possibility is to simply ignore the change, and apply the original plan $\widetilde{x}$ to the new situation $a = \widetilde{a} + \Delta a$.
- This plan is, in general, not optimal for the new situation.
- The actually optimal plan is $x^{\rm opt}$ for which

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

- In comparison with the optimal plan, we lose the amount $L_0 \stackrel{\rm def}{=} u(x^{\rm opt}, \widetilde{a} + \Delta a) - u(\widetilde{x}, \widetilde{a} + \Delta a)$.
- Why cannot we just find the optimal solution for the new situation?
- Optimization is NP-hard, so, it is not possible to find the exact optimum in reasonable time.

5. Options (cont-d)

- What we can do is:
  - try to use some feasible algorithm – e.g., solving a system of linear equations,
  - to modify the plan $\widetilde{x}$ into $\widetilde{x}+\Delta x$.
- Due to NP-hardness, this feasibly modified plan is, in general, not optimal.
- We hope that the resulting loss $L_1 \stackrel{\rm def}{=} u(x^{\rm opt},\widetilde{a}+\Delta a)-u(\widetilde{x}+\Delta x,\widetilde{a}+\Delta a)$ is much smaller than L_0.
- In this paper, we show that indeed $L_1 \ll L_0$; so:
  - even if L_0 is so large that the original plan is worthless,
  - the modified plan may leads to a reasonably small loss $L_1 \ll L_0$.
- This explains Eisenhower's observation.

6. 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 L_0 in Taylor series and keep only terms linear and quadratic in Δx:

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

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

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

- By definition, the action x^{opt} maximizes $u(x, \widetilde{a} + \Delta a)$.
- Thus, we have $\dfrac{\partial u}{\partial x_i}(x^{\text{opt}}, \widetilde{a} + \Delta a) = 0$.

7. Estimating L_0 (cont-d)

- So, the above expression for L_0 takes the simplified form
$$L_0 = \frac{1}{2}\cdot\sum_{i=1}^{n}\sum_{i'=1}^{n}\frac{\partial^2 u}{\partial x_i \partial x_{i'}}(x^{\rm opt},\widetilde{a}+\Delta a)\cdot\Delta x_i^{\rm opt}\cdot\Delta x_{i'}^{\rm opt}+o((\Delta a)^2).$$

- $\Delta x_i^{\rm opt}$ can be estimated from the condition:
$$\frac{\partial u}{\partial x_i}(x^{\rm opt},\widetilde{a}+\Delta a) = \frac{\partial u}{\partial x_i}(\widetilde{x}+\Delta x^{\rm opt},\widetilde{a}+\Delta) = 0.$$

- For $a=\widetilde{a}$, u is max when $x=\widetilde{x}$, so $\dfrac{\partial u}{\partial x_i}(\widetilde{x},\widetilde{a}) = 0$.

- Expanding the equation in Taylor series in Δx_i and Δa_j and taking $\dfrac{\partial u}{\partial x_i}(\widetilde{x},\widetilde{a}) = 0$ into account, we get:

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

8. Estimating L_0 (final)

- Thus, the first approximation Δx_i to the values Δx_i^{opt} satisfies a system of linear equations:

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

- 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 for L_0, we conclude that L_0 is quadratic in Δa_j:

$$L_0 = \sum_{j=1}^{m} \sum_{j'=1}^{m} k_{jj'} \cdot \Delta a_j \cdot \Delta a_{j'} + o((\Delta a)^2) \text{ for some } k_{jj'}.$$

9. Estimating L_1

- The 1st approximation Δx to the difference $\Delta x^{\rm opt}$ can be obtained by solving a system of linear equations.
- How much do we lose if we use $x^{\rm lin}=\widetilde{x}+\Delta x$?
- Here, $\Delta x^{\rm opt}=\Delta x+\delta x$, where δx is of 2nd order in Δx and Δa: $\delta x=O((\Delta a)^2)$.
- The loss L_1 of using $x^{\rm lin}=x^{\rm opt}-\delta x$ instead of $x^{\rm opt}$ is:

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

- If we expand this expression in δx, we get:

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

10. Estimating L_1 (cont-d)

- Since $x^{\rm opt}$ is the action that, for $a=\widetilde a+\Delta a$, maximizes utility, we get $\dfrac{\partial u}{\partial x_i}(x^{\rm opt},\widetilde a+\Delta a)=0.$
- 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*:

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

11. 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 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:

$$(1)\ \varepsilon < L_1,\ (2)\ L_1 \le \varepsilon \le L_0, \text{ and } (3)\ L_0 < \varepsilon.$$

- In the 1st case, even the modified action does not help.
- In the 3rd case, the change in the situation is so small that it is Ok to use the original plan $\widetilde{x}$.

12. Conclusions (cont-d)

- 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.

13. 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 from Prudential Foundation.