Why Inverse Layers in Pavement? Why Zipper Fracking? Why Interleaving in Education? A General Explanation

Edgar Daniel Rodriguez Velasquez[1,2], Aaron Velasco[3]
Olga Kosheleva[4], and Vladik Kreinovich[5]

[1,3,4,5]Departments of [1]Civil Engineering,
[3]Earth, Environment, and Resource Sciences,
[4]Teacher Education, [5]Computer Science
University of Texas at El Paso, El Paso, Texas 79968, USA
edrodriguezvelasquez@miners.utep.edu, aavelasco@utep.edu,
olgak@utep.edu, vladik@utep.edu
[2]Department of Civil Engineering, Universidad de Piura in Peru (UDEP),
Piura, Peru, edgar.rodriguez@udep.pe

1. Formulation of the Problem

- In several application areas, there appears a similar empirical phenomenon.
- In each of these areas, this phenomenon is difficult to explain.
- In this talk, we provide a general explanation for this phenomenon.
- Let us list the examples of this phenomenon.

2. Pavement engineering

- Road pavement must be strong enough to sustain the traffic loads.
- To strengthen the pavement, usually, the pavement is formed by the following layers.
- First, on top of the soil, we place compacted granular material; this is called the *sub-base*.
- On top of the sub-base, we place granular material strengthened with cement; this layer is called the *base*.
- Finally, the top layer is the granular material strengthened by adding the liquid asphalt; this layer is called the *asphalt concrete layer*.

3. Pavement engineering (cont-d)

- In this arrangement, the strength of the pavement comes largely from the two top layers: the asphalt concrete layer and the base.
- Empirical evidence shows that in many cases:
  - the inverse layer structure, where the base and sub-base are switched,
  - so that the two strong layers are separated by a weaker sub-base layer,
  - leads to better pavement performance.

4. Fracking

- Traditional methods of extracting oil and gas leave a significant portion of them behind.
- They were also unable to extract oil and gas that were concentrated in small amounts around the area.
- To extract this oil and gas, practitioners use the process called *fracking*.
- High-pressure liquid is injected into the underground location:
  - cracking the rocks and thus,
  - providing the path for low-density oil and gas to move to the surface.
- Usually, several pipes are used to pump the liquid.

5. Fracking (cont-d)

- Empirically, it turned out that the best performance happens:
  - not when all the pipes are active at the same time,
  - but when there is always a significant distance between the active pipes.
- One way to maintain this distance – known as *zipper fracking* – is to activate, e.g., every other pipe, interchanging:
  - activations of pipes 1, 3, 5, etc., with
  - activating the intermediate pipes 2, 4, 6, etc.
- This particular technique is known as *Texas two-step*.

6. Education

- In education, best learning results are achieved when there is a pause between two (or more) periods when some topic is studied.
- This pedagogical practice is known as *interleaving*.
- Several studies show that interleaving enhances different types of learning:
  - learning to play basketball,
  - learning art,
  - learning mathematics,
  - training and re-training medical doctors.

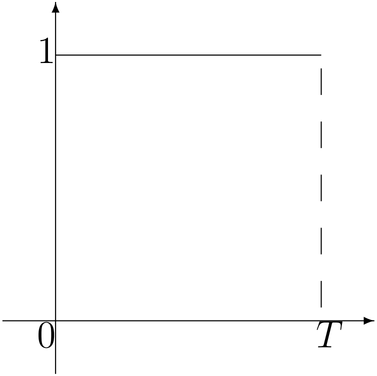
7. Towards an Explanation

- What is the ideal situation?
- The ideal pavement would mean that all layers are strong.
- The ideal fracking would mean that all the pipes are active all the time.
- The ideal study process would mean that we study all the time.
- So, a natural way to compare the quality of different strategies is to see which ones are closer to this ideal case.
- In general, we have a certain range; this can be the range that describes:
  - strength as a function of depth,
  - study intensity as a function of time, etc.
- From the mathematical viewpoint, we can always change the starting point to be 0.

8. Towards an Explanation (cont-d)

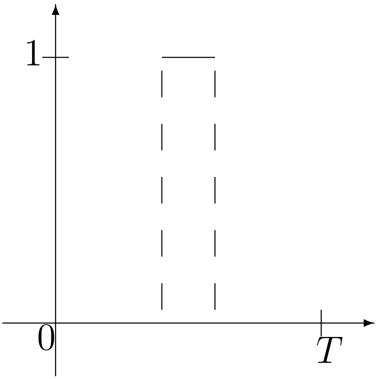
- For example, for studying, we can measure time starting with the moment when we started the whole study process.
- In this case, the range will take the form $[0, T]$ for some $T > 0$.
- So, for simplicity, let us assume that this range has the form $[0, T]$.
- Ideally, we should have full intensity at all points from this range:
  - we should have full strength at all depth,
  - we should have full study intensity at all moments of time, etc.
- From the mathematical viewpoint, we can re-scale intensity by taking this level as a new unit for measuring intensity.
- After this re-scaling, the value of the high level of intensity will be 1.
- So, the ideal case (I) is described by a function that takes the value 1 on the whole interval $[0, T]$.

9. Ideal Case



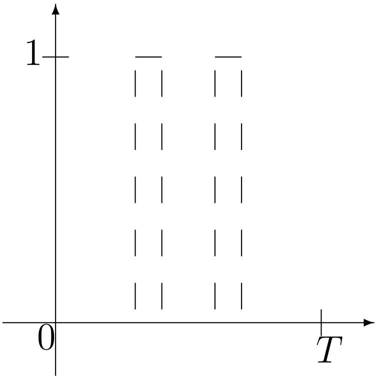
10. The ideal case is not realistic

- The problem is that in all the above applications, the ideal case is not realistic.
- In practice, we can have full strength only over a small portion of this range, a portion of overall size ε.
- We can have this strength portion concentrated on a connected (C) subrange.
- This is the case, e.g., of the traditional pavement.



11. The ideal case is not realistic (cont-d)

- Alternatively, we can divide this portion into two (or more) disconnected (D) subranges.



- In both cases, the value of intensity:
  - is equal to 1 on a small part of the range, and
  - is equal to 0 for all other values from the range.

12. Main idea behind our explanation

- In all the above examples, the performance was better for the disconnected subranges.
- We will explain this by proving that, in some reasonable sense,
  - the graph D corresponding to the disconnected portion is indeed closer to the graph I of the ideal case
  - than the graph corresponding to the connected portion C, i.e., that $d(D, I) < d(C, I)$.
- In order to prove this, let us recall what is the natural way to describe distance $d(A, B)$ between two graphs A and B.
- From the mathematical viewpoints, graphs are sets in a plane.
- So, to be able to describe distance between graphs, let us recall how to describe distance $d(A, B)$ between sets A and B.

13. How to define the distance $d(A, B)$ between two sets A and B: reminder

- Let us start with the simplest case, when both sets are 1-element sets, i.e., when $A = \{a\}$ and $B = \{b\}$ for some points a and b.
- We assume that for two points a and b, distance $d(a, b)$ is already defined.
- In this case, it is reasonable to define $d(A, B) = d(\{a\}, \{b\}) \stackrel{\text{def}}{=} d(a, b)$.
- A natural idea is to use Euclidean distance here:

$$d((x, y), (x', y')) = \sqrt{(x - x')^2 + (y - y')^2}.$$

- Instead, we can use a more general ℓ^p-metric for some $p \geq 1$:

$$d((x, y), (x', y')) = (|x - x'|^p + |y - y'|^p)^{1/p}.$$

- It is worth mentioning that our result remains valid whichever value $p \geq 1$ we select.

14. How to define the distance $d(A, B)$ (cont-d)

- A slightly more complex case is when only one of the sets is a one-point set, e.g., $A = \{a\}$.
- In this case, it makes sense to define the distance $d(\{a\}, B)$ in such a way that this distance is 0 when $a \in B$.
- A reasonable idea is to take

$$d(A, B) = d(\{a\}, B) \stackrel{\text{def}}{=} \inf_{b \in B} d(a, b).$$

- Finally, let us consider the general case, when both sets A and B may contain more than one point.
- In line with the general definition of a metric, we would like to have $d(A, B) = 0$ if and only if A and B coincide, i.e., if and only if:
  - every element the set A is also an element of the set B, and
  - every element of the set B is also an element of the set A.

15. How to define the distance $d(A, B)$ (cont-d)

- In other words, for us to declare that $d(A, B) = 0$:
  - we must have $d(\{a\}, B) = 0$ for all $a \in A$, and
  - we must have $d(\{b\}, A)$ for all $b \in B$.
- The usual way to achieve this purpose is – similarly to how we defined $d(\{a\}, B)$ – to define $d(A, B)$ as the largest of all these values.
- The resulting "worst-case" expression $d_w(A, B)$ is known as the *Hausdoff distance*:

$$d_w(A, B) \stackrel{\text{def}}{=} \max \left(\sup_{a \in A} d(\{a\}, B), \sup_{b \in B} d(\{b\}, A) \right).$$

- In general, the worst case is not always the most adequate description.
- For example, suppose that we have the set B almost equal to A, but with a very tiny additional part which is far away from A.
- In this case, the worst-case distance is huge, but in reality, the sets A and B are almost the same.

16. How to define the distance $d(A, B)$ (cont-d)

- To better capture the intuitive idea of distance between two sets, it is reasonable to consider:
  - not the *worst-case* values of $d(\{a\}, B)$ and $d(\{b\}, A)$, but
  - their *average* values:

$$d_a(A, B) \stackrel{\text{def}}{=} \frac{1}{2} \cdot \frac{\int_A d(\{a\}, B)\, da}{\mu(A)} + \frac{1}{2} \cdot \frac{\int_B d(\{b\}, A)\, db}{\mu(B)}.$$

- Let us see what these two definitions $d_w(A, B)$ and $d_a(A, B)$ say about the relation between our graphs I, C, and D.

17. What are the values $d_w(A, B)$ and $d_a(A, B)$ in our case

- Both worst-case and average-case definitions are based on the values $d(\{a\}, B)$ and $d(\{b\}, A)$.
- So, to compute the distances between the corresponding graphs, let us first find the values $d(\{a\}, B)$ and $d(\{b\}, A)$ for our case.
- Without losing generality, let us denote one of the graphs C or D by A, and the ideal graph I by B.
- Let us first consider the values $d(\{a\}, B) = d(\{a\}, I)$.
- Here, for points $a \in A$ corresponding to the portion of overall length ε, the intensity is equal to 1.
- These points also belong to the graph I and thus, $d(\{a\}, I) = 0$.
- For all other points $a \in A$, the intensity is 0, i.e., this point has the form $(x, 0)$ for some $x \in [0, T]$.
- The set I is the straight line segment.

18. What are the values $d_w(A, B)$ and $d_a(A, B)$ (cont-d)

- So, the closest element to I is the projection of the point A on this straight line, i.e., the point $(x, 1)$.
- In this case, the shortest distance $d(\{a\}, I)$ from the point a and points $b \in I$ is equal to 1: $d(\{a\}, I) = 1$.
- So, we have $\sup\limits_{a \in A} d(\{a\}, I) = 1$ and

$$\frac{\int_A d(\{a\}, I)\, da}{\mu(A)} = \frac{0 \cdot \varepsilon + 1 \cdot (T - \varepsilon)}{T} = \frac{T - \varepsilon}{T}.$$

- It should be mentioned that these values are the same both:
  - for the connected portion C and
  - for the disconnected portion D.
- Indeed, these values only depend on the overall length of the portion.
- Let us now consider the values $d(\{b\}, A)$, when $b \in I$, i.e., when $b = (x, 1)$ for some $x \in [0, 1]$, and A is C or D.

19. What are the values $d_w(A, B)$ and $d_a(A, B)$ (cont-d)

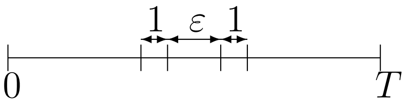
- By definition, $d(\{b\}, A)$ is the smallest of the values $d(a, b)$ when a is in the set A, i.e., when a is:
  - either in the portion – in which case $a = (x', 1)$ for some $x' \in [0, T]$,
  - or not in the portion – in which case $a = (x', 0)$ for some $x' \in [0, T]$.
- In the second case, the distance is at least 1 – and can always be made smaller than or equal to 1 if we take the point $(x, \cdot) \in A$.
- In the first case, the distance is equal to

$$d(a, b) = d((x, 1), (x', 1)) = |x - x'|.$$

- For points $b = (x, 1) \in I$ which are at most 1-close to the portion:
  - the shortest distance $d(\{b\}, A)$
  - is equal to the distance z between x and the portion.
- For all other points $b = (x, 1) \in I$, we have $d(\{b\}, A) = 1$.

20. What are the values $d_w(A, B)$ and $d_a(A, B)$ (cont-d)

- And herein lies the difference between the connected case C and the disconnected case D.
- In the connected case, we have:
  - one connected portion of length ε on which $d(\{b\}, A) = 0$, and
  - two nearby intervals for which $d(\{b\}, A) < 1$:



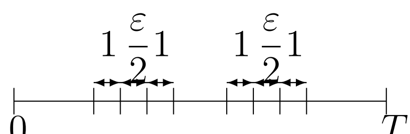
- Let us assume that:
  - that ε is sufficiently small, and
  - that the portion is sufficiently separated from the endpoints 0 and T of the range.

21. What are the values $d_w(A, B)$ and $d_a(A, B)$ (cont-d)

- Then, we have $\sup\limits_{b\in C} d(\{b\}, I) = 1$ and

$$\int_I d(\{b\}, C)\, db = 0\cdot\varepsilon + 2\int_0^1 z\, dz + (T-2-\varepsilon)\cdot 1 = 2\cdot\frac{1}{2} + T - 2 - \varepsilon = T - 1 - \varepsilon.$$

- Thus, $\dfrac{\int_I d(\{b\}, C)\, db}{\mu(I)} = \dfrac{T-1-\varepsilon}{T}$.
- In the disconnected case, we have:
  - two connected subranges (of length $\varepsilon/2$ each) on which $d(\{b\}, A) = 0$, and
  - two pairs of nearby intervals for which $d(\{b\}, A) < 1$.



22. What are the values $d_w(A,B)$ and $d_a(A,B)$ (cont-d)

- Let us also assume:
  - that ε is sufficiently small, and
  - that both subranges are sufficiently separated from each other and from the endpoints 0 and T of the range,
- Then, we have $\sup\limits_{b\in D} d(\{b\},I)=1$ and

$$\int_I d(\{b\},D)\,db=0\cdot\varepsilon+4\int_0^1 z\,dz+(T-2-\varepsilon)\cdot 1=4\cdot\frac{1}{2}+T-4-\varepsilon=T-2-\varepsilon.$$

- Thus $\dfrac{\int_I d(\{b\},C)\,db}{\mu(I)}=\dfrac{T-2-\varepsilon}{T}.$

23. Comparison summarized

- By combining the above formulas, we conclude that

$$d_w(C, I) = d_w(D, I) = 1.$$

- Thus, if we only take into account the worst-case distance, then we cannot distinguish between the connected and disconnected cases.

- However, if we use a more adequate average distance, then, by combining the above formulas, we get

$$d_a(C, I) = \frac{1}{2} \cdot \left(\frac{T - \varepsilon}{T} + \frac{T - 1 - \varepsilon}{T} \right) = \frac{T - 1/2 - \varepsilon}{T},$$

$$d_a(D, I) = \frac{1}{2} \cdot \left(\frac{T - \varepsilon}{T} + \frac{T - 2 - \varepsilon}{T} \right) = \frac{T - 1 - \varepsilon}{T}.$$

- Here clearly, $d_a(D, I) < d_a(C, I)$.

24. Comparison summarized (cont-d)

- In other words, the disconnected situation is closer to the ideal case than the connected one.
- This explains why in all above cases, the disconnected approach indeed leads to better results.

25. Conclusions and Future Work

- This paper provides a mathematical explanation for a phenomenon that is observed in reality but runs against our intuition:
  - a "disconnected" control, when there is a pause between the two control stages or a gap between two control locations,
  - is often more effective than the "connected" control, with no pauses or gaps.
- This is true in education, when:
  - learning the material during two time intervals, with a pause in between,
  - is often more effective than when these two time intervals immediately follow each other.

26. Conclusions and Future Work (cont-d)

- This is true in pavement engineering, where:
  - it is more effective to place two strong layers at some distance from each other
  - rather than placing them next to each other – as it is usually done.
- This is true in fracking, where keeping a distance between two active pipes is more effective than activating two neighboring pipes.
- In this talk, we provide a general mathematical explanation for this phenomenon.
- Our explanation is that:
  - for a naturally defined distance between settings,
  - the discontinuous setting is closer to the ideal one than the continuous setting.

27. Conclusions and Future Work (cont-d)

- The generality of this explanation makes us conjecture that a similar discontinuous arrangement is worth trying in many other cases:
  - automatic control,
  - medical therapy,
  - influencing people, etc
- Our first-approximation mathematical model will hopefully provide a way to compare different approaches.

28. Acknowledgments

This work was supported in part by:

- National Science Foundation grants 1623190, HRD-1834620, HRD-2034030, and EAR-2225395;
- AT&T Fellowship in Information Technology;
- program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and
- a grant from the Hungarian National Research, Development and Innovation Office (NRDI).