Towards Analytical Techniques for Optimizing Knowledge Acquisition, Processing, Propagation, and Use in Cyberinfrastructure

L. Octavio Lerma
Computational Science Program
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
lolerma@episd.org

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

1. Introduction

- Knowledge-related processes are important: we rely on them when we drive, communicate, etc.
- Surprisingly, the very process of acquiring and propagating information is the least automated.
- At present, to decide on the best way to place sensors or propagate data, we mostly use numerical models.
- These models are very resource-consuming, rely on supercomputers, not ready for everyday applications.
- We therefore need analytical models – which would allow easier optimization and application.
- Developing such models is our main objective.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

2. Outline

- We describe analytical models for all stages of knowledge processing.
- We start with knowledge acquisition: optimal sensor placement for stationary and mobile sensors.
- We then deal with data and knowledge processing: how to best organize computing power and research teams.
- We deal with knowledge propagation and resulting knowledge enhancement; we analyze:
  - how early stages of idea propagation occur;
  - how to assess the initial knowledge level;
  - how to present the material and how to provide feedback.
- Finally, we analyze how knowledge is used.

Introduction

Outline

Sensor Placement: . . .

Data and Knowledge . . .

Knowledge . . .

Resulting Geometric . . .

How to Use the . . .

Future Work

3. Sensor Placement: Case Study

- Biological weapons are difficult and expensive to detect.
- Within a limited budget, we can afford a limited number of bio-weapon detector stations.
- It is therefore important to find the optimal locations for such stations.
- A natural idea is to place more detectors in the areas with more population.
- However, such a commonsense analysis does not tell us how many detectors to place where.
- To decide on the exact detector placement, we must formulate the problem in precise terms.

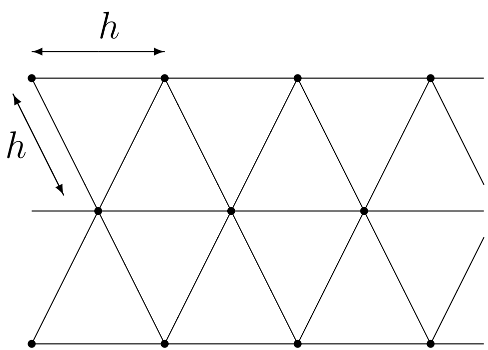
Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

4. Towards Precise Formulation of the Problem

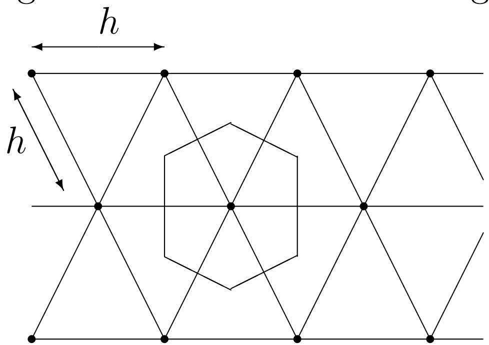
- The adversary's objective is to kill as many people as possible.
- Let $\rho(x)$ be a population density in the vicinity of the location x.
- Let N be the number of detectors that we can afford to place in the given territory.
- Let d_0 be the distance at which a station can detect an outbreak of a disease.
- Often, $d_0 = 0$ – we can only detect a disease when the sources of this disease reach the detecting station.
- We want to find $\rho_d(x)$ – the density of detector placement.
- We know that $\int \rho_d(x)\,dx = N$.

5. Optimal Placement of Sensors

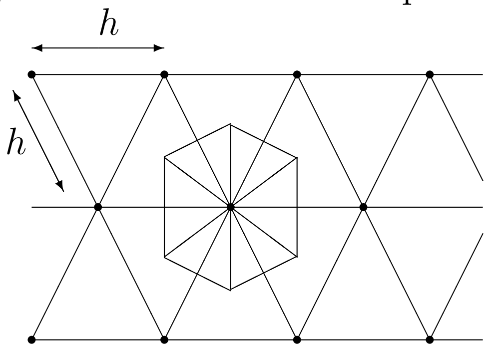
- We want to place the sensors in an area in such a way that
  - the largest distance D to a sensor
  - is as small as possible.
- It is known that the smallest such number is provided by an equilateral triangle grid:



For the equilateral triangle placement, points which are closest to a given detector forms a hexagonal area:



This hexagonal area consists of 6 equilateral triangles:



6. Optimal Placement of Sensors (cont-d)

- In each $\triangle$, the height $h/2$ is related to the side s by the formula $\dfrac{h}{2} = s \cdot \cos(60^\circ) = s \cdot \dfrac{\sqrt{3}}{2}$, hence $s = h \cdot \dfrac{\sqrt{3}}{3}$.
- Thus, the area A_t of each triangle is equal to

$$A_t = \frac{1}{2} \cdot s \cdot \frac{h}{2} = \frac{1}{2} \cdot \frac{\sqrt{3}}{3} \cdot \frac{1}{2} \cdot h^2 = \frac{\sqrt{3}}{12} \cdot h^2.$$

- So, the area A_s of the whole set is equal to 6 times the triangle area: $A_s = 6 \cdot A_t = \dfrac{\sqrt{3}}{2} \cdot h^2$.
- In a region of area A, there are $A \cdot \rho_d(x)$ sensors, they cover area $(A \cdot \rho_d(x)) \cdot A_s$.
- The condition $A = (A \cdot \rho_d(x)) \cdot A_s = (A \cdot \rho_d(x)) \cdot \dfrac{\sqrt{3}}{2} \cdot h^2$ implies that $h = \dfrac{c_0}{\sqrt{\rho_d(x)}}$, with $c_0 \stackrel{\text{def}}{=} \sqrt{\dfrac{2}{\sqrt{3}}}$.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

7. Estimating the Effect of Sensor Placement

- The adversary places the bio-weapon at a location which is the farthest away from the detectors.
- This way, it will take the longest time to be detected.
- For the grid placement, this location is at one of the vertices of the hexagonal zone.
- At these vertices, the distance from each neighboring detector is equal to $s = h \cdot \dfrac{\sqrt{3}}{3}$.
- By know that $h = \dfrac{c_0}{\sqrt{\rho_d(x)}}$, so $s = \dfrac{c_1}{\sqrt{\rho_d(x)}}$, with
$$c_1 = \frac{\sqrt{3}}{3} \cdot c_0 = \frac{\sqrt[4]{3} \cdot \sqrt{2}}{3}.$$
- Once the bio-weapon is placed, it starts spreading until it reaches the distance d_0 from the detector.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

8. Effect of Sensor Placement (cont-d)

- The bio-weapon is placed at a distance $s = \dfrac{c_1}{\sqrt{\rho_d(x)}}$ from the nearest sensor.
- Once the bio-weapon is placed, it starts spreading until it reaches the distance d_0 from the detector.
- In other words, it spreads for the distance $s - d_0$.
- During this spread, the disease covers the circle of radius $s - d_0$ and area $\pi \cdot (s - d_0)^2$.
- The number of affected people $n(x)$ is equal to:

$$n(x) = \pi \cdot (s - d_0)^2 \cdot \rho(x) = \pi \cdot \left(\frac{c_1}{\sqrt{\rho_d(x)}} - d_0\right)^2 \cdot \rho(x).$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

9. Precise Formulation of the Problem

- For each location x, the number of affected people $n(x)$ is equal to:

$$n(x) = \pi \cdot \left(\frac{c_1}{\sqrt{\rho_d(x)}} - d_0 \right)^2 \cdot \rho(x).$$

- The adversary will select a location x for which this number $n(x)$ is the largest possible:

$$n = \max_x \left(\pi \cdot \left(\frac{c_1}{\sqrt{\rho_d(x)}} - d_0 \right)^2 \cdot \rho(x) \right).$$

- Resulting problem:
  - given population density $\rho(x)$, detection distance d_0, and number of sensors N,
  - find a function $\rho_d(x)$ that minimizes the above expression n under the constraint $\int \rho_d(x)\,dx = N$.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

10. Main Lemma

- *Reminder:* we want to minimize the worst-case damage $n = \max_x n(x)$.
- *Lemma:* for the optimal sensor selection, $n(x) = \text{const}$.
- *Proof by contradiction:* let $n(x) < n$ for some x; then:
  - we can slightly increase the detector density at the locations where $n(x) = n$,
  - at the expense of slightly decreasing the location density at locations where $n(x) < n$;
  - as a result, the overall maximum $n = \max_x n(x)$ will decrease;
  - but we assumed that n is the smallest possible.
- *Thus:* $n(x) = \text{const}$; let us denote this constant by n_0.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

11. Towards the Solution of the Problem

- We have proved that $n(x) = \text{const} = n_0$, i.e., that

$$n_0 = \pi \cdot \left(\frac{c_1}{\sqrt{\rho_d(x)}} - d_0 \right)^2 \cdot \rho(x).$$

- Straightforward algebraic transformations lead to:

$$\rho_d(x) = \frac{2 \cdot \sqrt{3}}{9} \cdot \frac{1}{\left(d_0 + \frac{c_2}{\sqrt{\rho(x)}} \right)^2}.$$

- The value c_2 must be determined from the equation

$$\int \rho_d(x)\, dx = N.$$

- Thus, we arrive at the following solution.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

12. Solution

- *General case:* the optimal detector location is characterized by the detector density

$$\rho_d(x) = \frac{2\cdot\sqrt{3}}{9}\cdot\frac{1}{\left(d_0 + \frac{c_2}{\sqrt{\rho(x)}}\right)^2}.$$

- Here the parameter c_2 must be determined from the equation $\int \frac{2\cdot\sqrt{3}}{9}\cdot\frac{1}{\left(d_0 + \frac{c_2}{\sqrt{\rho(x)}}\right)^2}\,dx = N.$

- *Case of $d_0 = 0$:* in this case, the formula for $\rho_d(x)$ takes a simplified form $\rho_d(x) = C\cdot\rho(x)$ for some constant C.

- In this case, from the constraint, we get:

$$\rho_d(x) = \frac{N}{N_p}\cdot\rho(x),\ \text{ where } N_p \text{ is the total population.}$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

13. Towards More Relevant Objective Functions

- We assumed that the adversary wants to maximize the number $\int \rho(x)\,dx$ of people affected by the bio-weapon.
- The actual adversary's objective function may differ from this simplified objective function.
- For example, the adversary may take into account that different locations have different publicity potential.
- In this case, the adversary maximizes the weighted value $\int_A \widetilde{\rho}(x)\,dx$, where $\widetilde{\rho}(x) \stackrel{\text{def}}{=} w(x) \cdot \rho(x)$.
- Here, $w(x)$ is the importance of the location x.
- From the math. viewpoint, the problem is the same – w/"effective population density" $\widetilde{\rho}(x)$ instead of $\rho(x)$.
- Thus, if we know $w(x)$, we can find the optimal detector density $\rho_d(x)$ from the above formulas.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

14. How Temperatures etc. Change from One Spatial Location to Another: A Model

- Each environmental characteristic q changes from one spatial location to another.
- A large part of this change is unpredictable (i.e., random).
- A reasonable value to describe the random component of the difference $q(x)-q(x')$ is the variance

$$V(x,x') \stackrel{\text{def}}{=} E[((q(x)-E[q(x)])-(q(x')-E[q(x')]))^2].$$

- *Comment:* we assume that averages are equal.
- Locally, processes should not change much with shift $x \to x+s$: $V(x+s,x'+s)=V(x,x')$.
- For $s=-x'$, we get $V(x,x')=C(x-x')$ for

$$C(x) \stackrel{\text{def}}{=} V(x,0).$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

15. A Model (cont-d)

- In general, the further away the points x and x', the larger the difference $C(x - x')$.
- In the isotropic case, $C(x - x')$ depends only on the distance $D = |x - x'|^2 = (x_1 - x_1')^2 + (x_2 - x_2')^2$.
- It is reasonable to consider a scale-invariant dependence $C(x) = A \cdot D^\alpha$.
- In practice, we may have more changes in one direction and less change in another direction.
- E.g., 1 km in x is approximately the same change as 2 km in y.
- The change can also be mostly in some other direction, not just x- and y-directions.
- Thus, in general, in appropriate coordinates (u, v), we have $C = A \cdot D^\alpha$ for $D = (u - u')^2 + (v - v')^2$.

16. Model: Final Formulas

- In general, $C = A \cdot D^\alpha$, for $D = (u - u')^2 + (v - v')^2$ in appropriate coordinates (u, v).
- In the original coordinates x_1 and x_2, we get:

$$C(x - x') = A \cdot D^\alpha, \text{ where}$$

$$D = \sum_{i=1}^{2} \sum_{j=1}^{2} g_{ij} \cdot (x_i - x'_i) \cdot (x_j - x'_j) =$$

$$g_{11} \cdot (x_1 - x'_1)^2 + 2g_{12} \cdot (x_1 - x'_1) \cdot (x_2 - x'_2) + g_{22} \cdot (x_2 - x'_2)^2.$$

- From the computational viewpoint, we can include A into g_{ij} if we replace g_{ij} with $A^{1/\alpha} \cdot g_{ij}$, then

$$C(x - x') =$$

$$\left(g_{11} \cdot (x_1 - x'_1)^2 + 2g_{12} \cdot (x_1 - x'_1) \cdot (x_2 - x'_2) + g_{22} \cdot (x_2 - x'_2)^2\right)^\alpha$$

- We can use these formulas to find the optimal sensor locations.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

17. Optimal Use of Mobile Sensors: Case Study

- Remote areas of international borders are used by the adversaries: to smuggle drugs, to bring in weapons.
- It is therefore desirable to patrol the border, to minimize such actions.
- It is not possible to effectively man every single segment of the border.
- It is therefore necessary to rely on other types of surveillance.
- Unmanned Aerial Vehicles (UAVs):
  - from every location along the border, they provide an overview of a large area, and
  - they can move fast, w/o being slowed down by clogged roads or rough terrain.
- Question: what is the optimal trajectory for these UAVs?

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

18. How to Describe Possible UAV Patrolling Strategies

- Let us assume that the time between two consequent overflies is smaller the time needed to cross the border.
- Ideally, such a UAV can detect all adversaries.
- In reality, a fast flying UAV can miss the adversary.
- We need to minimize the effect of this miss.
- The faster the UAV goes, the less time it looks, the more probable that it will miss the adversary.
- Thus, the velocity $v(x)$ is very important.
- By a patrolling strategy, we will mean a f-n $v(x)$ describing how fast the UAV flies at different locations x.

19. Constraints on Possible Patrolling Strategies

1) The time between two consequent overflies should be smaller the time T needed to cross the border:

   – the time during which a UAV passes from the location x to the location $x+\Delta x$ is equal to $\Delta t = \dfrac{\Delta x}{v(x)}$;

   – thus, the overall flight time is equal to the sum of these times:
   $$T = \int \frac{dx}{v(x)}.$$

2) UAV has the largest possible velocity V, so we must have $v(x) \le V$ for all x.

It is convenient to use the value $s(x) \stackrel{\text{def}}{=} \dfrac{1}{v(x)}$ called *slowness*, so

$$T = \int s(x)\,dx; \quad s(x) \ge S \left(\stackrel{\text{def}}{=} \frac{1}{V} \right).$$

Introduction
Outline
Sensor Placement: ...
Data and Knowledge ...
Knowledge ...
Resulting Geometric ...
How to Use the ...
Future Work

20. Simplification of the Constraints

- Since $s(x) \geq S$, the value $s(x)$ can be represented as $S + \Delta s(x)$, where $\Delta s(x) \stackrel{\text{def}}{=} s(x) - S$.
- The new unknown function satisfies the simpler constraint $\Delta s(x) \geq 0$.
- In terms of $\Delta s(x)$, the requirement that the overall time be equal to T has a form $T = S \cdot L + \int \Delta s(x)\, dx$.
- This is equivalent to:

$$T_0 = \int \Delta s(x)\, dx, \text{ where:}$$

  - L is the total length of the piece of the border that we are defending, and
  - $T_0 \stackrel{\text{def}}{=} T - S \cdot L$.

Introduction
Outline
Sensor Placement: ...
Data and Knowledge ...
Knowledge ...
Resulting Geometric ...
How to Use the ...
Future Work

21. Detection at Crossing Point x

- Let h be the width of the border zone from which an adversary (A) is visible.
- Then, the UAV can potentially detect A during the time $h/v(x) = h \cdot s(x)$.
- So, the UAV takes $(h \cdot s(x))/\Delta t$ photos, where Δt is the time per photo.
- Let p_1 be the probability that one photo misses A.
- It is reasonable to assume that different detection errors are independent.
- Then, the probability $p(x)$ that A is not detected is $p_1^{(h \cdot s(x))/\Delta t}$, i.e., $p(x) = \exp(-k \cdot s(x))$, where:

$$k \stackrel{\text{def}}{=} \frac{2h}{\Delta t} \cdot |\ln(p_1)|.$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

22. Strategy Selected by the Adversary

- Let $w(x)$ denote the utility of the adversary succeeding in crossing the border at location x.
- Let us first assume that we know $w(x)$ for every x.
- According to decision theory, the adversary will select a location x with the largest expected utility

$$u(x) = p(x) \cdot w(x) = \exp(-k \cdot s(x)) \cdot w(x).$$

- Thus, for each slowness function $s(x)$, the adversary's gain $G(s)$ is equal to

$$G(s) = \max_x u(x) = \max_x \left[\exp(-k \cdot s(x)) \cdot w(x)\right].$$

- We need to select a strategy $s(x)$ for which the gain $G(s)$ is the smallest possible.

$$G(s) = \max_x u(x) = \max_x \left[\exp(-k \cdot s(x)) \cdot w(x)\right] \to \min_{s(x)}.$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

23. Towards an Optimal Strategy for Patrolling the Border

- Let x_m be the location at which the utility $u(x) = \exp(-k \cdot s(x)) \cdot w(x)$ attains its largest possible value.
- If we have a point x_0 s.t. $u(x_0) < u(x_m)$ and $s(x_0) > S$:
  - we can slightly decrease the slowness $s(x_0)$ at the vicinity of x_0 (i.e., go faster in this vicinity) and
  - use the resulting time to slow down (i.e., to go slower) at all locations x at which $u(x) = u(x_m)$.
- As a result, we slightly decrease the value

$$u(x_m) = \exp(-k \cdot s(x_m)) \cdot w(x_m).$$

- At x_0, we still have $u(x_0) < u(x_m)$.
- So, the overall gain $G(s)$ decreases.
- Thus, when the adversary's gain is minimized, we get

$$u(x) = u_0 = \text{const whenever } s(x) > S.$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

24. Towards an Optimal Strategy (cont-d)

- *Reminder:* for the optimal strategy,

  $$u(x) = w(x) \cdot \exp(-k \cdot s(x)) = u_0 \text{ whenever } s(x) > S.$$

- So, $\exp(-k \cdot s(x)) = \dfrac{u_0}{w(x)}$, hence

  $$s(x) = \frac{1}{k} \cdot (\ln(w(x)) - \ln(u_0)) \text{ and } \Delta s(x) = \frac{1}{k} \cdot \ln(w(x)) - \Delta_0.$$

- Here, $\Delta_0 \stackrel{\text{def}}{=} \dfrac{1}{k} \cdot \ln(u_0) - S$.
- When $s(x)$ gets to $s(x) = S$ and $\Delta s(x) = 0$, we get $\Delta s(x) = 0$.
- Thus, we conclude that

  $$\Delta s(x) = \max\left(\frac{1}{k} \cdot \ln(w(x)) - \Delta_0, 0\right).$$

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

25. An Optimal Strategy: Algorithm

- *Reminder:* for some Δ_0, the optimal strategy has the form

$$\Delta s(x) = \max\left(\frac{1}{k}\cdot \ln(w(x)) - \Delta_0, 0\right).$$

- *How to find Δ_0:* from the condition that

$$\int \Delta s(x)\,dx =$$

$$\int \max\left(\frac{1}{k}\cdot \ln(w(x)) - \Delta_0, 0\right)\,dx = T_0.$$

- *Easy to check:* the above integral monotonically decreases with Δ_0.

- *Conclusion:* we can use bisection to find the appropriate value Δ_0.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

26. Efficient Algorithms for Optimizing Sensor Use: Case Study of Security Problems

- In this section, we analyze the problem of designing efficient algorithms for optimizing resource allocations.
- Case study: protection of critical infrastructure from terrorist attacks, computer network security, etc.
- Previously known algorithms for optimal resource allocation required quadratic time $O(n^2)$.
- We develop new algorithms which require time

$$O(n \cdot \log(n)) \ll O(n^2).$$

- In important special cases, this algorithm runs even faster, in linear time $O(n)$.

27. Data and Knowledge Processing: How to Best Organize Computing and Human Resources

- Once the data is collected, we need to process this data.
- For processing, we need computing power, and we need human resources.
- In both cases, we need to come up with optimal resource allocation.
- In Section 3.1, we come up with the optimal allocation formulas for computing resources.
- In Section 3.2, we come up with the optimal allocation formulas for human resources.
- The corresponding mathematics is similar to the optimal distribution of sensors.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

28. Knowledge Propagation and Resulting Knowledge Enhancement

- Once we have transformed data into knowledge, we need to propagate this knowledge.
- For that, we first need to motivate people to learn the new knowledge.
- To ensure this, we analyze the process of knowledge propagation.
- On early stages, the number of knowledgeable people grows as a power law t^α.
- In Section 4.1, we provide a theoretical explanation for this empirical fact.
- Once a person is interested, we assess how much he/she knows, and how to best teach the material.
- This is covered in Sections 4.2-4.5.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

29. Assessing the Initial Knowledge Level

- Computers enable us to provide individualized learning, at a pace tailored to each student.
- In order to start the learning process, it is important to find out the current level of the student's knowledge.
- Usually, such placement tests use a sequence of N problems of increasing complexity.
- If a student is able to solve a problem, the system generates a more complex one.
- If a student cannot solve a problem, the system generates an easier one, etc.
- Once we find the exact level of student's knowledge, the actual learning starts.
- It is desirable to get to actual leaning as soon as possible, i.e., to minimize the # of placement problems.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

30. Bisection – Optimal Search Procedure

- At each stage, we have:
  - the largest level i at which a student can solve, &
  - the smallest level j at which s/he cannot.
- Initially, $i = 0$ (trivial), $j = N + 1$ (very tough).
- If $j = i + 1$, we found the student's level of knowledge.
- If $j > i + 1$, give a problem on level $m \stackrel{\text{def}}{=} (i + j)/2$:
  - if the student solved it, increase i to m;
  - else decrease j to m.
- In both cases, the interval $[i, j]$ is decreased by half.
- In s steps, we decrease the interval $[0, N + 1]$ to width $(N + 1) \cdot 2^{-s}$.
- In $s = \lceil \log_2(N + 1) \rceil$ steps, we get the interval of width ≤ 1, so the problem is solved.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

31. Need to Account for Discouragement

- Every time a student is unable to solve a problem, he/she gets discouraged.
- In bisection, a student whose level is 0 will get $\approx \log_2(N+1)$ negative feedbacks.
- For positive answers, the student simply gets tired.
- For negative answers, the student also gets stressed and frustrated.
- If we count an effect of a positive answer as one, then the effect of a negative answer is $w > 1$.
- The value w can be individually determined.
- We need a testing scheme that minimizes the worst-case overall effect.

32. Analysis of the Problem

- We have $x = N + 1$ possible levels of knowledge.
- Let $e(x)$ denote the smallest possible effect needed to find out the student's knowledge level.
- We ask a student to solve a problem of some level n.
- If s/he solved it (effect $= 1$), we have $x - n$ possible levels $n, \ldots, N$.
- The effect of finding this level is $e(x - n)$, so overall effect is $1 + e(x - n)$.
- If s/he didn't (effect w), his/her level is between 0 and n, so we need effect $e(n)$, with overall effect $w + e(n)$.
- Overall worst-case effect is $\max(1+e(x-n), w+e(n))$.
- In the optimal test, we select n for which this effect is the smallest, so $e(x) = \min\limits_{1 \le n < x} \max(1+e(x-n), w+e(n))$.

33. Resulting Algorithm

- For $x = 1$, i.e., for $N = 0$, we have $e(1) = 0$.
- We know that $e(x) = \min\limits_{1 \le n < x} \max(1 + e(x - n), w + e(n))$.
- We can use this formula to sequentially compute the values $e(2)$, $e(3)$, $\ldots$, $e(N + 1)$.
- We also compute the corresponding minimizing values $n(2)$, $n(3)$, $\ldots$, $n(N + 1)$.
- Initially, $i = 0$ and $j = N + 1$.
- At each iteration, we ask to solve a problem at level $m = i + n(j - i)$:
  - if the student succeeds, we replace i with m;
  - else we replace j with m.
- We stop when $j = i + 1$; this means that the student's level is i.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

34. Example 1: $N = 3$, $w = 3$

- Here, $e(1) = 0$.
- When $x = 2$, the only possible value for n is $n = 1$, so

$$e(2) = \min_{1 \le n < 2} \{\max\{1 + e(2 - n), 3 + e(n)\}\} =$$

$$\max\{1 + e(1), 3 + e(1)\} = \max\{1, 3\} = 3.$$

- Here, $e(2) = 3$, and $n(2) = 1$.
- To find $e(3)$, we must compare two different values $n = 1$ and $n = 2$:

$$e(3) = \min_{1 \le n < 3} \{\max\{1 + e(3 - n)), 3 + e(n)\}\} =$$

$$\min\{\max\{1+e(2), 3+e(1)\}, \max\{1+e(1), 3+e(2)\}\} =$$

$$\min\{\max\{4, 3\}, \max\{1, 6\}\} = \min\{4, 6\} = 4.$$

- Here, min is attained when $n = 1$, so $n(3) = 1$.

35. Example 1: $N = 3$, $w = 3$ (cont-d)

- To find $e(4)$, we must consider three possible values $n = 1$, $n = 2$, and $n = 3$, so

$$e(4) = \min_{1 \le n < 4} \{\max\{1 + e(4 - n), 3 + e(n))\}\} =$$

$$\min\{\max\{1 + e(3), 3 + e(1)\}, \max\{1 + e(2), 3 + e(2)\},$$

$$\max\{1 + e(1), 3 + e(3)\}\} =$$

$$\min\{\max\{5, 3\}, \max\{4, 6\}, \max\{1, 7\}\} =$$

$$\min\{5, 6, 7\} = 5.$$

- Here, min is attained when $n = 1$, so $n(4) = 1$.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

36. Example 1: Resulting Procedure

- First, $i = 0$ and $j = 4$, so we ask a student to solve a problem at level $i + n(j - i) = 0 + n(4) = 1$.
- If the student fails level 1, his/her level is 0.
- If s/he succeeds at level 1, we set $i = 1$, and we assign a problem of level $1 + n(3) = 2$.
- If the student fails level 2, his/her level is 1.
- If s/he succeeds at level 2, we set $i = 2$, and we assign a problem of level $2 + n(3) = 3$.
- If the student fails level 3, his/her level is 2.
- If s/he succeeds at level 3, his/her level is 3.
- We can see that this is the most cautious scheme, when each student has at most one negative experience.

37. Example 2: $N = 3$ and $w = 1.5$

- We take $e(1) = 0$.
- When $x = 2$, then

$$e(2) = \min_{1 \le n < 2} \{\max\{1 + e(2-n), 3 + e(n)\}\} =$$

$$\max\{1 + e(1), 1.5 + e(1)\} = \max\{1, 1.5\} = 1.5.$$

- Here, $e(2) = 1.5$, and $n(2) = 1$.
- To find $e(3)$, we must compare two different values $n = 1$ and $n = 2$:

$$e(3) = \min_{1 \le n < 3} \{\max\{1 + e(3-n)), 1.5 + e(n)\}\} =$$

$$\min\{\max\{1+e(2), 1.5+e(1)\}, \max\{1+e(1), 1.5+e(2)\}\} =$$

$$\min\{\max\{2.5, 1.5\}, \max\{1, 3\}\} = \min\{2.5, 3\} = 2.5.$$

- Here, min is attained when $n = 1$, so $n(3) = 1$.

38. Example 2: $N = 3$ and $w = 1.5$ (cont-d)

- To find $e(4)$, we must consider three possible values $n = 1$, $n = 2$, and $n = 3$, so

$$e(4) = \min_{1 \le n < 4} \{\max\{1 + e(4-n), 1.5 + e(n))\}\} =$$

$$\min\{\max\{1+e(3), 1.5+e(1)\}, \max\{1+e(2), 1.5+e(2)\},$$

$$\max\{1 + e(1), 1.5 + e(3)\}\} =$$

$$\min\{\max\{3.5, 1.5\}, \max\{2.5, 3\}, \max\{1, 4\}\} =$$

$$\min\{3.5, 3, 4\} = 3.$$

- Here, min is attained when $n = 2$, so $n(4) = 2$.

39. Example 2: Resulting Procedure

- First, $i = 0$ and $j = 4$, so we ask a student to solve a problem at level $i + n(j - i) = 0 + n(4) = 2$.
- If the student fails level 2, we set $j = 2$, and we assign a problem of level $0 + n(2) = 1$:
  - if the student fails level 1, his/her level is 0;
  - if s/he succeeds at level 1, his/her level is 1.
- If s/he succeeds at level 2, we set $i = 2$, and we assign a problem at level $2 + n(2) = 3$:
  - if the student fails level 3, his/her level is 2;
  - if s/he succeeds at level 3, his/her level is 3.
- We can see that in this case, the optimal testing scheme is bisection.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

40. A Faster Algorithm May Be Needed

- For each n from 1 to N, we need to compare n different values.
- So, the total number of computational steps is proportional to $1 + 2 + \ldots + N = O(N^2)$.
- When N is large, N^2 may be too large.
- In some applications, the computation of the optimal testing scheme may takes too long.
- For this case, we have developed a faster algorithm for producing a testing scheme.
- The disadvantage of this algorithm is that it is only asymptotically optimal.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

41. A Faster Algorithm for Generating an Asymptotically Optimal Testing Scheme

- First, we find the real number $\alpha \in [0,1]$ for which $\alpha + \alpha^w = 1$.
- This value α can be obtained, e.g., by applying bisection to the equation $\alpha + \alpha^w = 1$.
- At each iteration, once we know bounds i and j, we ask the student to solve a problem at the level
$$m = \lfloor \alpha \cdot i + (1 - \alpha) \cdot j \rfloor.$$
- This algorithm is similar to bisection, expect that bisection corresponds to $\alpha = 0.5$.
- This makes sense, since for $w = 1$, the equation for α takes the form $2\alpha = 1$, hence $\alpha = 0.5$.
- For $w = 2$, the solution to the equation $\alpha + \alpha^2 = 1$ is the well-known golden ratio $\alpha = \dfrac{\sqrt{5} - 1}{2} \approx 0.618$.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

42. Towards Optimal Teaching

- One of the main objectives of a course – calculus, physics, etc. – is to help students understand its main concepts.
- Of course, it is also desirable that the students learn the corresponding methods and algorithms.
- However, understanding is the primary goal.
- If a student does not remember a formula by heart, she can look it up.
- However:
  - if a student does not have a good understanding of what, for example, is a derivative,
  - then even if this student remembers some formulas, he will not be able to decide which formula to apply.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

43. How to Gauge Student Understanding

- To properly gauge student's understanding, several disciplines have developed *concept inventories*.
- These are sets of important basic concepts and questions testing the students' understanding.
- The first such Force Concept Inventory (FCI) was developed to gauge the students' understanding of forces.
- A student's degree of understanding is measured by the percentage of the questions that are answered correctly.
- The class's degree of understanding is measured by averaging the students' degrees.
- An ideal situation is when everyone has a perfect 100% understanding; in this case, the average score is 100%.
- In practice, the average score is smaller than 100%.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

44. How to Compare Different Teaching Techniques

- We can measure the average score μ_0 before the class and the average score μ_f after the class.
- Ideally, the whole difference $100 - \mu_0$ disappears, i.e., the students' score goes from μ_0 to $\mu_f = 100$.
- In practice, of course, the students' gain $\mu_f - \mu_0$ is somewhat smaller than the ideal gain $100 - \mu_0$.
- It is reasonable to measure the success of a teaching method by which portion of the ideal gain is covered:

$$g \stackrel{\text{def}}{=} \frac{\mu_f - \mu_0}{100 - \mu_0}.$$

Introduction
Outline
Sensor Placement: ...
Data and Knowledge ...
Knowledge ...
Resulting Geometric ...
How to Use the ...
Future Work

45. Empirical Results

- It turns out that the gain g does not depend on the initial level μ_0, on the textbook used, or on the teacher.
- Only one factor determines the value g: the absence or presence of immediate feedback.
- In traditionally taught classes,
  - where the students get their major feedback only after their first midterm exam,
  - the average gain is $g \approx 0.23$.
- For the classes with an immediate feedback, the average gain is twice larger: $g \approx 0.48$.
- In this talk, we provide a possible geometric explanation for this doubling of the learning rate.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

46. Why Geometry

- Learning means changing the state of a student.
- At each moment of time, the state can be described by the scores $x_1, \ldots, x_n$ on different tests.
- Each such state can be naturally represented as a point $(x_1, \ldots, x_n)$ in the n-dimensional space.
- In the starting state S, the student does not know the material.
- The desired state D describes the situation when a student has the desired knowledge.
- When a student learns, the student's state of knowledge changes continuously.
- It forms a (continuous) trajectory γ which starts at the starting state S and ends up at the desired state D.

Introduction
Outline
Sensor Placement: ...
Data and Knowledge ...
Knowledge ...
Resulting Geometric ...
How to Use the ...
Future Work

47. First Simplifying Assumption: All Students Learn at the Same Rate

- Some students learn faster, others learn slower.
- The above empirical fact, however, is not about their *individual* learning rates.
- It is about the *average* rates of student learning, averaged over all kinds of students.
- From this viewpoint, it makes sense to assume that all the students have the same average learning rate.
- In geometric terms, this means that the leaning time is proportional to the length of the corresponding curve γ.
- We thus need to show that learning trajectories corr. to immediate feedback are, on average, twice shorter.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

48. Second Simplifying Assumption: the Shape of the Learning Trajectories

- At first, a student has misconceptions about physics or calculus, which lead him in a wrong direction.
- We can thus assume that at first, a student moves in a random direction.
- After the feedback, the student corrects his/her trajectory.
- In the case of immediate feedback, this correction comes right away, so the students goes in the right direction.
- In the traditional learning, with a midterm correction:
  - a student first follows a straight line of length $d/2$ which goes in a random direction,
  - and then takes a straight line to the midpoint M.
- Then, a student goes from M to the destination D.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

49. 3rd Simplifying Assumption: 1-D State Space

- We can think of different numerical characteristics describing different aspects of student knowledge.
- In practice, to characterize the student's knowledge, we use a single number – the overall grade for the course.
- It is therefore reasonable to assume that the state of a student is characterized by only one parameter x_1.
- In case of immediate feedback, the learning trajectory has length d.
- To make a comparison, we must estimate the length of a trajectory corresponding to the traditional learning.
- This trajectory consists of two similar parts: connecting S and M and connecting M and D.
- To estimate the total average length, we can thus estimate the average length from S to M and double it.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

50. Analysis: Case of Traditional Leaning

- A student initially goes either in the correct direction or in the opposite (wrong) direction.
- Randomly means that both directions occur with equal probability $1/2$.
- If the student moves in the right direction, she gets exactly into the desired midpoint M.
- In this case, the length of the S-to-M part of the trajectory is exactly $d/2$.
- If the student starts in the wrong direction, he ends up at a point at distance $d/2$ – on the wrong side of S.
- Getting back to M then means first going back to S and then going from S to M.
- The overall length of this trajectory is thus $3d/2$.

51. Resulting Geometric Explanation

- Here:
  - with probability $1/2$, the length is $d/2$;
  - with probability $1/2$, the length is $3d/2$.
- So, the average length of the S-to-M part of the learning trajectory is equal to

$$\frac{1}{2} \cdot \frac{d}{2} + \frac{1}{2} \cdot \frac{3d}{2} = d.$$

- The average length of the whole trajectory is double that, i.e., $2d$.
- This average length is twice larger than the length d corresponding to immediate feedback.
- This explains why immediate feedback makes learning, on average, twice faster.

52. How to Use the Resulting Knowledge: Case Study

- How can we use the acquired knowledge?
- In many practical situations, we have a well-defined problem, with a clear well-formulated objective.
- Such problems are typical in engineering:
  - we want a bridge which can withstand a given load,
  - we want a car with a given fuel efficiency, etc.
- However, in many practical situations, it is important to also take into account subjective user preferences.
- This subjective aspect of decision making is known as *Kansei engineering*.

53. Need to Select Designs

- Different people have different preferences.
- Thus, to satisfy customers, we must produce several different designs:
  - a car company produces cars of several different designs,
  - a furniture company produces chairs of several different designs, etc.
- The creation of each new design is often very expensive and time-consuming.
- As a result, the number of new designs is usually limited.
- Once we know what customers want and how many designs we can afford, how to select these designs?

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

54. Results

- In Chapter 5, we come up with analytical formulas for optimal design selection.
- These formulas are similar to formulas of optimal sensor allocation.
- Each design can be characterized by an n-dimensional vector $x = (x_1, \ldots, x_n)$.
- Each user has his/her own ideal design x.
- For many users, we have a density $\rho_u(x)$ corr. to different designs.
- We provide an analytical formula describing the optimal design density $\rho_m(x)$.

55. Future Work: Main Theoretical Activity

- Recently, a seminal book appeared *Social Physics* by A. Pentland from MIT.
- This book describes the successful results of using models to enhance knowledge propagation.
- This book describes many well-justified results.
- It also describes interesting empirical observations for which no theoretical explanations are available.
- Our plan is to look into these observations and results and see if some of them can be theoretically explained.

56. Future Work: Auxiliary Theoretical Activity

- Petland's book views knowledge propagation from the viewpoint of mathematical optimization.
- We thus try to find the best values of the parameters of the corresponding knowledge propagation process.
- In this model, all the decisions are centralized.
- Educational practice shows that often, efficiency can be drastically improved by decentralization.
- Specifically, we allow teachers – and students – to select different ways of propagating knowledge.
- There have been several empirical studies of this phenomenon.
- We plan to look for a theoretical explanation for the known empirical results.

Introduction
Outline
Sensor Placement: . . .
Data and Knowledge . . .
Knowledge . . .
Resulting Geometric . . .
How to Use the . . .
Future Work

57. Future Work: Practical Applications

- The ultimate goal of the theoretical research is to enhance actual knowledge propagation.
- As part of our research, we have already developed some practical recommendations.
- We plan to test these recommendations on the actual processes of teaching and knowledge propagation.
- In particular, we plan to test them on cyberinfrastructure-related data acquisition, processing, and propagation.
- These applications are what motivated our research.
- We thus hope that our recommendations will be useful for cyberinfrastructure-related applications.

58. Acknowledgments

- My deep gratitude to my committee: Drs. V. Kreinovich, D. Pennington, C. Tweedie, and S. Starks.
- I also want to thank all the faculty, staff, and students of the Computational Science program.
- My special thanks to Drs. M. Argaez, A. Gates, M.-Y. Leung, A. Pownuk, L. Velazquez, and to C. Davis.
- Last but not the least, my thanks and my love to my family:
  - to my sons Sam and Joseph,
  - to my parents Hortensia and Antonio,
  - to my sister Martha, and
  - to my brothers Antonio, Victor, and Javier.

Introduction
Outline
Sensor Placement: ...
Data and Knowledge ...
Knowledge ...
Resulting Geometric ...
How to Use the ...
Future Work