1-2015

Why Right-Brain Cultures Are More Flexible: A Possible Explanation of Yu. Manin's Observation

Olga Kosheleva
University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
University of Texas at El Paso, vladik@utep.edu

Follow this and additional works at: http://digitalcommons.utep.edu/cs_techrep

Part of the Mathematics Commons

Comments:
Published in International Mathematical Forum, 2015, Vol. 10, No. 4, pp. 175-180.

Recommended Citation
http://digitalcommons.utep.edu/cs_techrep/896

This Article is brought to you for free and open access by the Department of Computer Science at DigitalCommons@UTEP. It has been accepted for inclusion in Departmental Technical Reports (CS) by an authorized administrator of DigitalCommons@UTEP. For more information, please contact lweber@utep.edu.
Why Right-Brain Cultures Are More Flexible: A Possible Explanation of Yu. Manin’s Observation

Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract
Yuri Manin, a renowned mathematician, observed that it is much easier for a person raised in a right-brain culture to adjust to the left-brain environment than vice versa. In this paper, we provide a possible explanation for this phenomenon.

1 Manin’s Observation: Formulation of the Problem

Manin’s observation: formulation of the problem. In his talk [?], Yuri Manin, a renowned mathematician, made several observations about cultures that place more emphasis on left-brain, logical, discrete reasoning (e.g., traditional Western cultures) and cultures that place more emphasis on right-brain, intuitive, continuous reasoning (e.g., traditional Chinese culture and other Oriental cultures).

One of his observations is that it is much easier for a person raised in the traditional right-brain culture to adopt to a left-brain one than vice versa. For example:

• many students and researchers who are originally from China blossom in the Western academic institutions, but

• for a Western researcher it is much more difficult to fully adjust to a Chinese academic institution.

What we plan to do. In this paper, we provide a possible mathematical explanation for this phenomenon.
2 Formalization of the Problem

Need for a simplified model. To understand the corresponding adjustment phenomena, let us consider the simplest possible mathematical model for such adjustment. For that, we need to formalize what we mean by right-brain (continuous) and left-brain (discrete) cultures, and what we mean by adjustment.

Right-brain (continuous) and left-brain (discrete) cultures: a simplified description. In mathematics, the most natural notion of continuity is the continuity of a function. From this viewpoint, we will model:

- knowledge representations corresponding to right-brain (continuous) cultures by continuous functions, and
- knowledge representations corresponding to left-brain (discrete) cultures by discrete (piece-wise constant) functions.

We need to select finite-parametric families of functions. At any given moment of time, we can only store finitely many parameters, so we have to restrict ourselves to finite-parametric families of functions. How do we select these families?

Functional dependencies $y = f(x)$ are ubiquitous in nature. For example, for a given body:

- the force $y$ is a function of acceleration $x$ (second Newton’s law),
- the voltage $y$ is a function of the current $x$ (Ohm’s law),
- the acceleration $y$ caused by the Sun’s gravity is a function of the distance $x$,

etc. In all these cases, we can observe the values $y_i = f(x_i)$ of the desired function $f(x)$ corresponding to different inputs $x_1 < \ldots < x_n$:

- forces corresponding to different accelerations,
- voltages corresponding to different values of the current, etc.

It is therefore reasonable to consider functions generated by this information, i.e., by finitely many values $x_1, \ldots, x_n$ and the corresponding values $y_1, \ldots, y_n$.

Resulting families. In the continuous case, for every two tuples $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$, we need to consider a continuous function $f(x)$ for which $f(x_i) = y_i$. The simplest such function is a piece-wise linear function that connects the corresponding points, i.e., a function for which:

- $f(x) = y_1$ for $x \leq x_1$,
- $f(x) = y_n$ for $x \geq x_n$, and
\[ f(x) = y_i + \frac{x - x_i}{x_{i+1} - x_i} \text{ when } x_i \leq x \leq x_{i+1}. \]

In the discrete case, the simplest possible idea is to use a piece-wise constant function whose value \( f(x) \) at each point \( x \) coincides with the value \( y_i \) at the nearest point \( x_i \). In this case,

- \( f(x) = y_1 \) for \( x < \frac{x_1 + x_2}{2} \),
- \( f(x) = y_n \) for \( x \geq \frac{x_{n-1} + x_n}{2} \), and
- \( f(x) = y_i \) for \( \frac{x_{i-1} + x_i}{2} \leq x < \frac{x_i + x_{i+1}}{2} \).

**What does adjustment mean in these terms?** Since we are talking about functions, a natural meaning of adjustment is approximation:

- for a right-brain (continuous) person to adjust to the left-brain (discrete) culture means that we try to approximate a discrete function by continuous ones;
- similarly, for a left-brain (discrete) person to adjust to the right-brain (continuous) culture means that we try to approximate a continuous function by discrete ones.

As a numerical measure of approximation quality, we can use the Least Squares difference – the one that is mostly used in data processing. In our case, this means that, in effect, as a measure of difference between two functions \( f(x) \) and \( g(x) \) we consider the quantity \( \int (f(x) - g(x))^2 \, dx \). For this integral to be finite, we need to limit ourselves to a finite interval \( [L, U] \).

Now, we are ready to formulate the problem in precise terms.

### 3 Definitions and the Main Result

**Definition 1.** Let \( L < U \) be real numbers, and let \( n \) be a positive integer.

- **By an \( n \)-parametric continuous function**, we mean a function that, for some tuples \( L \leq x_1 < x_2 < \ldots < x_n \leq U \) and \( (y_1, \ldots, y_n) \), has the form
  
  - \( f(x) = y_1 \) for \( x \leq x_1 \),
  - \( f(x) = y_n \) for \( x \geq x_n \), and
  - \( f(x) = y_i + \frac{x - x_i}{x_{i+1} - x_i} \text{ when } x_i \leq x \leq x_{i+1} \).

- **By an \( n \)-parametric discrete function**, we mean a function that, for some tuples \( L \leq x_1 < x_2 < \ldots < x_n \leq U \) and \( (y_1, \ldots, y_n) \), has the form
  
  - \( f(x) = y_1 \) for \( x < \frac{x_1 + x_2}{2} \),
  - \( f(x) = y_n \) for \( x \geq \frac{x_{n-1} + x_n}{2} \), and
  - \( f(x) = y_i \) for \( \frac{x_{i-1} + x_i}{2} \leq x < \frac{x_i + x_{i+1}}{2} \).
\[ f(x) = y_n \text{ for } x \geq \frac{x_{n-1} + x_n}{2}, \text{ and } \]

\[ f(x) = y_i \text{ for } \frac{x_{i-1} + x_i}{2} \leq x < \frac{x_i + x_{i+1}}{2}. \]

**Definition 2.** By a distance between two functions \( f(x) \) and \( g(x) \), we mean the value \( d(f, g) \equiv \int_U (f(x) - g(x))^2 \, dx \). For every real number \( \varepsilon > 0 \), we say that \( f \) and \( g \) are \( \varepsilon \)-close if \( d(f, g) \leq \varepsilon \).

**Proposition 1.** For every \( n \) and for every \( \varepsilon > 0 \), for every \( n \)-parametric continuous function \( f(x) \), there is a \( 2n \)-parametric discrete function \( g(x) \) which is \( \varepsilon \)-close to \( f(x) \).

**Proposition 2.** For every \( n \) and \( N > n \), for every non-constant \( n \)-parametric continuous function \( f(x) \), there exists a real value \( \varepsilon > 0 \) such that no \( N \)-parametric discrete function is \( \varepsilon \)-close to \( f(x) \).

**Conclusion.** So, while a discrete function can be approximated by continuous functions with any possible accuracy, there is always a limit with which a discrete function can approximate a continuous one. In other words, it is much easier for continuous functions to approximate discrete ones than vice versa, it is much easier for a continuous function to adjust to a discrete family than vice versa. This is exactly what we tried to explain.

### 4 Proofs

**Proof of Proposition 1.** Let \( f(x) \) be a continuous function corresponding to the values \( x_1 < \ldots < x_n \) and \( y_1, \ldots, y_n \). To create an approximating discrete function, we approximate each jump by a fast-increasing linear function.

Specifically, let us select a small \( \delta > 0 \), and consider \( 2(n-1) \) values \( x^-_1 < x^+_2 < x^-_2 < \ldots < x^-_{n-1} < x^+_{n-1} \), where \( x^-_i = \frac{x_{i+1} + x_i}{2} - \delta \) and \( x^+_i = \frac{x_{i+1} + x_i}{2} + \delta \). As the corresponding \( y \)-values, we take \( y^-_i \equiv y_i \) and \( y^+_i \equiv y_{i+1} \).

One can check that the resulting \( 2n \)-parametric continuous function \( g(x) \) differs from the original discrete function \( f(x) \) only on \( n-1 \) intervals \([x^-_i, x^+_i]\) of length \( 2\delta \) each – and thus, of total length \( 2(n-1) \cdot \delta \). On these intervals, the values of both functions \( f(x) \) and \( g(x) \) are bounded from below by \( y \equiv \min(y_1, \ldots, y_n) \) and from above by \( Y \equiv \max(y_1, \ldots, y_n) \). Thus,

\[ |f(x) - g(x)| \leq Y - y, \]

hence \((f(x) - g(x))^2 \leq (Y - y)^2\), and therefore, \( d(f, g) \leq (Y - y)^2 \cdot 2(n-1) \cdot \delta \).

So, when \( \delta \to 0 \), we get \( d(f, g) \to 0 \). The statement is proven.
Proof of Proposition 2. Since the \( n \)-parametric continuous function \( f(x) \) is not a constant, there exists an \( i \) for which \( y_i \neq y_{i+1} \). On the corresponding interval \([x_i, x_{i+1}]\), the function \( f(x) \) is linear, with a non-zero slope \( k \equiv \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \).

Any \( N \)-parametric discrete function \( g(x) \) is based on \( N \) pairs \((x_i', y_i')\) that divide the whole interval \([L, U]\) into \( N \) subintervals on each of which the function \( g(x) \) is constant. In particular, some of these intervals cover the whole interval \([x_i, x_{i+1}]\). Thus, the interval \([x_i, x_{i+1}]\) is divided into \( \leq N \) parts of each of which \( g(x) \) is a constant. The sum of the lengths of these parts is \( x_{i+1} - x_i \), thus at least one of these parts must have the length \( \ell \) which is at least \( h \equiv \frac{x_{i+1} - x_i}{N} \) – else, if all their lengths are smaller, their sum would be smaller than \( x_{i+1} - x_i \).

On this part of length \( \ell \), a linear function \( f(x) \) with slope \( k \) is approximated by a constant. The distance \( d(f, g) = \int_{L}^{U} (f(x) - g(x))^2 \, dx \) between \( f(x) \) and \( g(x) \) is greater than or equal to the integral \( \int (f(x) - g(x))^2 \, dx \) limited to this part. One can easily check that for a linear function, the smallest \( \ell^2 \)-distance from a constant is when this constant is equal to a midpoint between the two extreme points of the linear function. On this part, the corresponding integral is thus equal to

\[
\int_{-\ell/2}^{\ell/2} (k \cdot x)^2 \, dx = 2 \int_{0}^{\ell/2} k^2 \cdot x^2 \, dx = 2k^2 \cdot \frac{1}{3} \cdot \left( \frac{\ell}{2} \right)^3.
\]

Since the distance \( d(f, g) \) is larger than or equal to this integral and \( \ell \geq h \), we have

\[
d(f, g) \geq 2k^2 \cdot \frac{1}{3} \cdot \left( \frac{h}{2} \right)^3.
\]

Any positive real number which is smaller than the right-hand side of this inequality can thus be taken as the desired value \( \varepsilon > 0 \). The proposition is proven.

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.

References