Why Bump Reward Function Works Well In Training Insulin Delivery Systems

Lehel Dénes-Fazakas\textsuperscript{1}, László Szilágyi\textsuperscript{1,2}, Gyorgy Eigner\textsuperscript{1}, Olga Kosheleva\textsuperscript{3}, Vladik Kreinovich\textsuperscript{3}, and Nguyen Hoang Phuong\textsuperscript{4}

\textsuperscript{1}\'Obuda University, Bécsi Street 96/B, Budapest H-1034, Hungary
denes-fazakas.lehel@uni-obuda.hu, eigner.gyorgy@uni-obuda.hu

\textsuperscript{2}Sapientia University, Tg. Mureş, Romania, lalo@ms.sapientia.ro

\textsuperscript{3}University of Texas at El Paso, 500 W. University, El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu

\textsuperscript{4}Artificial Intelligence Division, Information Technology Faculty
Thang Long University, Nghiem Xuan Yem Road
Hoang Mai District, Hanoi, Vietnam, nhphuong2008@gmail.com
1. What is diabetes

- All living creatures need energy to function.
- To many cells in a human body, energy comes from glucose that is delivered to these cells by the blood flow.
- The absorption of glucose into the cells is regulated by a special hormone called insulin.
- Sometimes, the body does not produce enough insulin – the illness known as diabetes.
- Diabetes hinders the ability of cells to get energy and can lead to life-threatening situations.
2. Need for insulin injections

- To avoid dangerous situations, a natural idea is to inject insulin into the body when the insulin level becomes dangerously low.

- We can detect that, since in this case:
  - the cells do not absorb the glucose and thus,
  - the blood glucose level becomes too high.

- In healthy patients, the body itself decides how much insulin is needed.

- With diabetes, there is no automatic biologic regulation.

- So, at present, the patients themselves decide, based on some general recommendations:
  - when to inject insulin and
  - how much to inject.

- The effectiveness of these general recommendations is different for different patients.
3. Need for insulin injections (cont-d)

- It is therefore desirable to have automatic systems individually trained to become maximally effective for each patient.
- Such systems are indeed being actively developed, trained, and tested all over the world.
4. Empirical fact: bump reward function works the best

- The purpose of the system is to keep the patient’s blood sugar level $x$ within the desired interval $[x, \bar{x}]$.
- For training the automatic insulin delivery system, we can, in principle, use different reward functions.
- Researchers compared the results of using different reward functions.
- They found out that the most effective is the so-called *bump* reward function that is:
  - outside the desired interval, equal to 0, and
  - for values $x$ within this interval, equal to
    \[
    b(x) = \exp \left( -\frac{c}{(x - \underline{x}) \cdot (\bar{x} - x)} \right)
    \]
- A natural question is: why the bump functions works best?
- In this talk, we provide a possible explanation for the effectiveness of the bump function.
5. **What is a natural reward function?**

- We want the patient to feel healthy.
- At each moment of time, the only information that we have about the patient is the patient’s blood glucose level $x$.
- Based on this level, we can only determine the probability $p(x)$ that the patient feels healthy.
- This probability is what we want to maximize, i.e., that should be our reward function.
- Clearly, when the value $x$ is outside the given interval, something is wrong, so the corresponding probability is 0 (or close to 0).
- So, to find an appropriate reward function, we need to find the probabilities $p(x)$ corresponding to values $x$ from the given interval.
6. We need to select probabilities based on partial information

- In many practical situations, probabilities are determined experimentally, as corresponding frequencies.

- However, in our case, we do not have enough statistics, so we need to select the probability distribution based on whatever information we have.

- For this purpose, let us recall how, in general, a probability distribution is determined based on partial information.

- In many practical situations, we only have partial information about probabilities.

- For example:
  - we may know that there are two possible situations, but
  - we have no information which of the two situations is more probable.
7. We need to select probabilities based on partial information (cont-d)

- In such situations, a reasonable idea is to assign equal probability to both situations.
- Similarly:
  - if we have $n$ possible situations, and we have no reason to believe that one of them is more probable,
  - a reasonable idea is to assign, to all of them, equal probability $1/n$.
- This natural idea is known as *Laplace Indeterminacy Principle*.
- This principle can be described in a slightly different way.
- If we have two alternatives, we have an uncertainty, in which:
  - to determine which is a correct one,
  - we need to ask one binary ("yes"-"no") question.
8. We need to select probabilities based on partial information (cont-d)

- If we have $2^n$ alternatives, then we need $n$ binary questions to uniquely determine the alternative.

- When we select equal probabilities, the average number of questions needed to determine the situation remains the same.

- However, if we selected unequal probabilities, then, on average, the number of questions becomes smaller.

- For example:
  - if we assign probability 1 to one of the alternatives and 0 to all others,
  - we need 0 questions to find the alternative – it is the one whose probability is 1.

- So, in this case we kind of cheat, we insert artificial certainty where there was none.
9. We need to select probabilities based on partial information (cont-d)

- Similarly:
  - if we have partial information about probabilities,
  - i.e., if there is a whole set of probability distributions that is consistent with available information,
  - then a reasonable idea is not to cheat, not to add artificial certainty, but to preserve the original uncertainty.

- In the discrete case, a natural measure of uncertainty is:
  - the average number of binary questions that is needed
  - to uniquely determine the alternative.

- In the continuous case, a similar natural measure is:
  - the average number of binary question that is needed
  - to determine the unknown value $x$ with a given accuracy $\varepsilon > 0$. 
10. **We need to select probabilities based on partial information** (cont-d)

- In both cases, there are distributions for which this average number of questions is smaller.
- However, selecting them would be artificially adding certainty.
- What we need is:
  - the distribution that best reflects the original uncertainty,
  - i.e., for which the average number of questions is as large as possible.
- It turns out that the average of question is described by Shannon’s entropy $S = - \int f(x) \cdot \ln(f(x)) \, dx$.
- Here $f(x)$ is the corresponding probability density function.
11. We need to select probabilities based on partial information (cont-d)

- Thus, a reasonable idea is to select, from each class of probability distributions, the distribution with the largest possible entropy.

- This *Maximum Entropy* approach has indeed led to many successful applications.
12. Let us apply this general idea to our case: first idea

- We have a class of probability distributions located on the interval $[x, \bar{x}]$.
- We do not make any assumptions about the distribution.
- So, the only constraint on the probability density function is that the overall probability is 1: $\int f(x) \, dx = 1$.
- To maximize entropy under this constraint, it is natural to use Lagrange multiplier method.
- This method reduces a constraint optimization problem to an equivalent unconstrained optimization problem.
- In our case, we need to maximizing the following function:

$$- \int f(x) \cdot \ln(f(x)) \, dx + \lambda \cdot \left( \int f(x) \, dx - 1 \right).$$

- Here, the value $\lambda$ – known as Lagrange multiplier – is determined by the condition that the optimal function $f(x)$ satisfies the constraint.
13. First idea (cont-d)

- The solution to this unconstrained optimization problem can be obtained:
  - by using the known fact from calculus
  - that the maximum of an expression is attained when all its derivatives are equal to 0.

- Differentiating the expression (with respect to each unknown \( f(x) \)) and equating the derivative to 0, we conclude that
  \[- \ln(f(x)) - 1 + \lambda = 0.\]

- Thus \( f(x) = \exp(\lambda - 1) = \text{const}. \)

- So, we get a uniform distribution on the desired interval.

- This is in perfect accordance with the above argument – that used Laplace Indeterminacy Principle.
14. Limitations of the first idea

- From the mathematical viewpoint, this is reasonable.
- However, from the viewpoint of our problem, it is not reasonable at all.
- Indeed, for the uniform distribution, the probability of being healthy is exactly the same:
  - whether we are in the middle of the desired interval
  - or we are close to one of its endpoints.
- However, in practice:
  - if the value of the blood glucose level start getting closer to the threshold,
  - this should be a sign to be alarmed – so the probability of being healthy should be smaller close to the endpoints.
15. Limitations of the first idea (cont-d)

- This means that we cannot get a reasonable distribution if we do not impose any additional constraints.
- We need to impose some additional constraints if we want a reasonable result.
16. Second idea

- We need to add constraints, and constraints reflect partial information that we have.
- What do we know about the probability distribution?
- We rarely know individual characteristic, but often, from observations, we know averages.
- So, a seemingly natural idea is to add a constraint that we know the average value $\tilde{x}$ of the quantity $x$: $\int x \cdot f(x) \, dx = \tilde{x}$.
- Let us apply the same Lagrange multiplier method to the problem of maximizing entropy under new constraint.
- We arrive at the problem of optimizing the following expression:

$$- \int f(x) \ln(f(x)) \, dx + \lambda_1 \left( \int f(x) \, dx - 1 \right) + \lambda_2 \left( \int x \cdot f(x) \, dx - \tilde{x} \right).$$

- Equating derivatives to 0 leads to $-\ln(f(x)) = 1 + \lambda_1 + \lambda_2 \cdot x = 0$, so $f(x) = \exp((\lambda_1 - 1) + \lambda_2 \cdot x)$. 
17. Limitations of the second idea

- The resulting formula has the same limitations as the first idea:
  - we want the probability of healthiness to tend to 0 as approach the endpoints,
  - but this is not happening here.
18. How can we modify the second idea?

- Let us take into account that the same physical quantity can be described by different numerical values.
- First, we can select a different measuring unit.
- For example, the height of 2 m becomes 200 if we use centimeters.
- Second, we can select a different starting point.
- For example, the current year 2023 can become year 2014 in Ethiopian calendar that uses a different starting date.
- Finally, many quantities are ratios.
19. How can we modify the second idea (cont-d)

- For example, blood glucose level is the ratio of the amount of glucose in blood to the corresponding amount of blood.

- In such cases, we can reverse the ratio and also get a meaningful description of the same quantity; for example:
  - we can have velocity $v = \frac{d}{t}$ – which is the ratio of distance to time – and we can have slowness $\frac{1}{v} = \frac{t}{d}$;
  - we can have resistance $R = \frac{V}{I}$ – which is the ratio of voltage to current – and we can have conductivity $\frac{1}{R} = I \cdot V$, etc.
20. First try

- If we simply change the measuring unit or the starting point, the situation does not change.

- Fixing the mean value of the re-scaled quantity $k \cdot x$ or $x + x_0$ is equivalent to fixing the mean value of the quantity $x$ itself.

- What if we reverse the formula for the blood sugar level and consider the mean value of this reverse $\int \frac{1}{x} \cdot f(x) \, dx = \tilde{r}$.

- Then the corresponding constraint optimization leads to

$$ - \ln(f(x)) - 1 + \lambda_1 + \lambda_2 \cdot \frac{1}{x} = 0. $$

- So $f(x) = \exp \left( (\lambda_1 - 1) + \frac{\lambda_2}{x} \right)$.

- This is still not exactly we want.
21. Second try leads to the desired explanation

- What if we take into account both the possibility of taking a reverse \textit{and} the probability of changing the starting points.

- It is reasonable to use both endpoints $x$ and $\bar{x}$ as starting points; thus:
  - we get the re-scaled values $x - \underline{x}$ and $x - \bar{x}$ (or, better, $\bar{x} - x$, to keep the values non-negative), and
  - reversing these re-scaled values leads to the following two constraints: $\int \frac{1}{x - \underline{x}} \cdot f(x) \, dx = \tilde{r}_-, \int \frac{1}{\bar{x} - x} \cdot f(x) \, dx = \tilde{r}_+$.

- Maximizing the entropy under these constraints leads to the following unconstrained optimization problem:
  \begin{align*}
  - \int f(x) \cdot \ln(f(x)) \, dx + \lambda _1 \cdot \left( \int f(x) - 1 \right) &+ \\
  \lambda_- \cdot \left( \int \frac{1}{x - \underline{x}} \cdot f(x) \, dx - \tilde{r}_- \right) &+ \lambda_+ \cdot \left( \int \frac{1}{\bar{x} - x} \cdot f(x) \, dx - \tilde{r}_+ \right).
  \end{align*}
22. Second try leads to the desired explanation (cont-d)

- For this expression, equating its derivatives to 0 leads to
  \[ \ln(f(x)) = -1 + \lambda_1 + \frac{\lambda_-}{x - \overline{x}} + \frac{\lambda_+}{\overline{x} - x}. \]

- Here:
  - the value \( \lambda_- \) reflects the importance of the left endpoint of the desired interval,
  - while the value \( \lambda_+ \) reflects the uncertainty of the right endpoint.

- Both endpoints are important:
  - going beyond each of these two endpoints can be life-threatening, and
  - we have no reason to assume that one of the endpoints is more important.

- Thus, in line with the Laplace Indeterminacy Principle, it makes sense to assume that these two values are equal: \( \lambda_- = \lambda_+ \).
• In this case, the above formula takes the form

\[ \ln(f(x)) = \text{const} + \frac{\text{const}}{(x - \overline{x}) \cdot (\overline{x} - x)}. \]

• So, we get exactly the bump function expression for the probability values \( f(x) \)!

• Thus, we have indeed explained the effectiveness of the bump reward function.

• The explanation is that this function naturally follows from first principles.
24. 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).