

## ADJUSTMENT OF A POSITIVE INTEGER PARAMETER UNKNOWN TO AN INTERVAL WITH CONSTANT BOUNDARIES BASED ON EXPERT ESTIMATIONS WHOSE AVERAGE-LIKE VALUE IS UPPER-LIMITED TO THE PARAMETER

*Adjustment of an unknown parameter relating to expert procedures and estimations is represented. The parameter is a positive integer. Boundaries of an interval enclosing the parameter's value are fixed constants. The goal is to develop a parameter adjustment algorithm based on expert estimations, wherein an average-like value of those estimations must be not greater than the parameter. First off, a simpler adjustment algorithm is developed. This is the hard adjustment, wherein an inequality with the parameter is checked and a subsequent mean dichotomy is fulfilled. If the inequality, in which the average-like value must be not greater than the parameter, turns false then it implies that the right end has been taken too small. And then the simplest correction becomes dragging it to the top right. The soft adjustment algorithm corrects it smarter dragging it to some value which was previously set to be the current parameter's value. Because of the average-like value of expert estimations is upper-limited to the parameter, dichotomy is fulfilled by the prioritized shift to the right. Both algorithms provide an early stop condition, when the average-like value is approximately equal to the parameter's value and it is not greater than the parameter for a row of cases. The hard adjustment algorithm should be applied in fields of poor knowledge or with less proficient experts. The soft adjustment algorithm fits strong experience and highly proficient experts.*

*Keywords: positive integer parameter adjustment, expert procedure, expert estimations, tolerance.*

В. В. РОМАНЮК  
Хмельницький національний університет

### ПРИПАСУВАННЯ НЕВІДОМОГО ДОДАТНОГО ЦІЛОЧИСЕЛЬНОГО ПАРАМЕТРА У ФОРМІ ІНТЕРВАЛУ З ПОСТІЙНИМИ ГРАНИЦЯМИ НА ОСНОВІ ЕКСПЕРТНИХ ОЦІНОК, ЧИС ПОДІБНЕ ДО СЕРЕДНЬОГО ЗНАЧЕННЯ ОБМЕЖЕНО ДАНИМ ПАРАМЕТРОМ ЗВЕРХУ

*Представляється припасування невідомого параметра, що відноситься до експертних процедур й оцінювань. Параметр є додатним цілим числом. Границі інтервалу, котрий включає значення параметра, є відомими константами. Мета полягає у розробці алгоритму припасування параметра на основі експертних оцінювань, де подібне до середнього значення цих оцінювань має бути не більшим за даний параметр. Перш за все, розробляється алгоритм простішого припасування. Це — жорстке припасування, де перевіряється одна нерівність з даним параметром, та в результаті виконується серединна дихотомія. Якщо ця нерівність, в якій подібне до середнього значення має бути не більшим за параметр, виявляється хибною, то це означає, що правий кінець був узятий занадто малим. І тоді відбувається найпростіша корекція зі збільшенням відтягуванням праворуч. Алгоритм м'якого припасування корегує більш витончено з відтягуванням до деякого значення, яке попередньо вже було поточним значенням параметра. Через те, що подібне до середнього значення експертних оцінювань обмежено зверху даним параметром, дихотомія виконується за пріоритетним зсувом праворуч. Обидва алгоритми надають умову ранньої зупинки, коли подібне до середнього значення приблизно дорівнює значенню параметра і воно не більше за цей параметр протягом декількох випадків поспіль. Алгоритм жорсткого припасування має застосовуватись у галузях з низьким рівнем знань або з менш кваліфікованими експертами. Алгоритм м'якого припасування відповідає ґрунтовному досвіду та висококваліфікованим експертам.*

*Ключові слова: припасування додатного цілочисельного параметра, експертна процедура, експертні оцінки, допуск.*

### Adjustment of positive integers

Almost every field of study has its integer parameters whose values are not calculated or deduced directly, but are assigned by experience. Just as experience is building up, the integer parameter's value is assigned more accurate. Then the parameter unknown to an interval with constant boundaries, if any, is said that it is adjusted. The most relevant fields concerning much adjustment of positive integers are computer science [1, 2] and socio-environmental statistics [3, 4], which both still require huge amount of expert experience. In particular, statistical approximators like convolutional neural networks [5, 6] have a lot of integer hyperparameters [7, 8]. Their lower boundaries are obviously fixed (e. g., minimal number of filters, dimension filter shape, max-pooling size), and upper boundaries are limited as well. Adjustment of those positive integers is similar to that in other fields: starting off the average-near integers (like setting the hidden layer neuron number in two-layer perceptron [9]), they then are refined based on new problems solved with older integers. The refinement is fulfilled owing to data of expert procedures or test statistics. Commonly, the adjustment comes from one of the boundaries towards the second boundary.

### Known approaches to adjust positive integers

The key approach to adjusting a positive integer is its successive refinement on the basis of experiments [6, 7, 10, 11], tests [2, 6, 7, 9, 12, 13], measurements [14, 15], expert procedures [3, 4, 16, 17], etc. The pivot point is an average-like value drawn from those procedures compared to its previous state (value). Due to that the parameter is integer and it is unknown to an interval with constant boundaries, the set  $G$  of all its possible values is finite. This warrants another adjustment approach, which contemplates random selection of the parameter's values from the set  $G$ , whereupon the most successful values are retained constituting the next set  $G_* \subset G$  by  $G_* \cap G = G_*$  [3, 6, 8, 9, 13, 17]. Thus, the initial set  $G$  is narrowed step by step until a single value is retained.

Drawbacks of these known approaches are their weak theoretical substantiation and specificity focus. There is no unique routine of adjustment. Moreover, due to that basic experimental procedure often changes (for instance, datasets for training and testing convolutional neural networks), the pivot point is drawn vaguer and unclear [3, 9].

### Goal and bullets to be accomplished in order to meet it

The stated drawbacks induce the goal which is to develop an algorithm of adjusting a positive integer parameter unknown to an interval with constant boundaries. For definiteness, the adjustment shall be based on expert estimations. Besides, an average-like value of those estimations shall be not greater than the parameter.

In order to meet the goal, there are four bullets to be accomplished:

1. A simple adjustment algorithm is developed. Actually, it should rather be the simplest one, to be a starting routine which can be modified easier if necessary. The simple adjustment algorithm should regard long-term procedures and have an early stop condition when the pivot point comes stable enough.

2. The developed algorithm is modified so that it could be softer when the pivot point comes greater than the parameter, though previously it was less than or equal to the parameter. The simpler adjustment algorithm should be a particular case of the modified algorithm.

3. Application of the suggested adjustment is briefly and clearly described. This is an appendix to the adjustment algorithms. The application must clarify when simpler and softer algorithms should be preferably used.

4. The suggested adjustment is discussed for revealing its merits and demerits. The demerits, if any, will help in outlining further work more pragmatically, to possibly improve the adjustment concerning specifically intricate fields and their problems.

### Hard adjustment of the integer parameter

Denote the parameter's value by  $\gamma$ ,  $\gamma \in \mathbb{N}$ . Let  $\gamma_{\max}$  and  $\gamma_{\min}$  be the upper and lower boundaries, respectively. By the convention,  $\gamma_{\max} \in \mathbb{N}$  and  $\gamma_{\min} \in \mathbb{N}$ . Basically, an average-like value  $a$  of expert estimations is upper-limited to the parameter, and so the inequality

$$a \leq \gamma \quad (1)$$

is conceptual and crucial. Value  $a \in [\gamma_{\min}; \gamma_{\max}]$  and it does not have to be an integer, because expert estimations are usually diverse and manipulations over them (even if they are integer numbers) do not always give integers.

Throughout expert procedures,  $\gamma$  is rendered by  $a$ . Let the number of expert procedures be limited to  $Q_{\max}$ ,  $Q_{\max} \in \mathbb{N} \setminus \{1\}$ . And let tolerate a row of  $Q^*$  cases when  $a$  is approximately equal to  $\gamma$ , but inequality (1) holds, where  $Q^* \in \mathbb{N} \setminus \{1\}$  and  $Q^* \leq Q_{\max}$ .

Before the first procedure starts, the parameter's value is set in the roughest manner being the simplest ever. This is just the mean of  $\gamma_{\max}$  and  $\gamma_{\min}$ . If

$$\frac{\gamma_{\max} + \gamma_{\min}}{2} \in \mathbb{N} \quad (2)$$

then the parameter's value is set to

$$\gamma = \frac{\gamma_{\max} + \gamma_{\min}}{2}. \quad (3)$$

Otherwise, if membership (2) is false,

$$\gamma = \frac{\gamma_{\max} + \gamma_{\min} + 1}{2}. \quad (4)$$

It is worth noting that a new point  $\gamma$  cannot be closer to  $a$  than to the obsolete value of  $\gamma$ , which is  $\gamma_{\max}$  (before the first procedure). The reason is that we cannot squeeze the uncertain interval too much to the left, or else inequality (1) may be violated. Let this reason be called the right-side priority.

Inequality (1) is checked always after the  $q$ -th procedure is executed and its value  $a$  based on experts' estimations is calculated,  $q = \overline{1, Q_{\max}}$ . If (1) is true then, similarly to (3) or (4), a new value of  $\gamma$  is found by the right-side priority. However, inequality

$$\gamma - a < 2 \quad (5)$$

along with inequality (1) means that the current value of  $\gamma$  is not changed (Figure 1). If (5) is false then it is

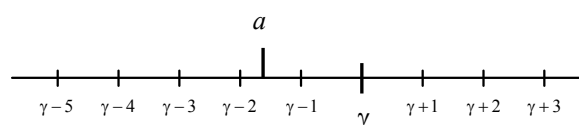


Figure 1. The case when both inequalities (1) and (5) are true, and the current value of  $\gamma$  is not changed due to  $\gamma-1$  is closer to  $a$  than to  $\gamma$ , so changeover to  $\gamma-1$  would have broken the right-side priority

checked on whether  $a$  is approximately equal to  $\gamma$  (or whether they are too close) using some tolerance  $\delta$ . If

$$\frac{\gamma - a}{\gamma_{\max} - \gamma_{\min}} < \delta \quad (6)$$

then, again,  $a$  is said to be approximately equal to  $\gamma$  (these points are too close), and the current value of  $\gamma$  is not changed. If (6) is false, then, due to the right-side priority, a new value of  $\gamma$  is found as the mean of the obsolete value of  $\gamma$  and the integer part of  $a$  increased by either 1 or 2. If

$$\frac{\gamma + \varphi(a) + 1}{2} \in \mathbb{N} \quad (7)$$

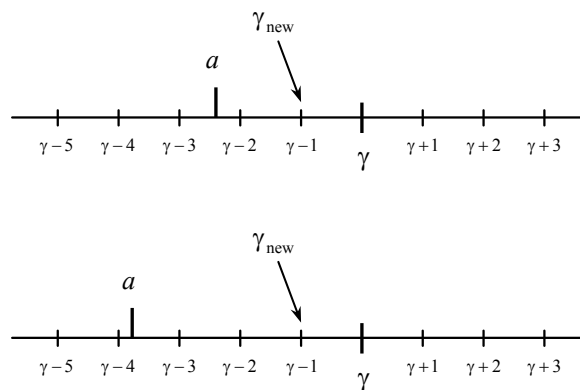
by the function  $\varphi(x)$  returning the integer part of real number  $x$ , then

$$\gamma_{\text{new}} = \frac{\gamma + \varphi(a) + 1}{2} \quad (8)$$

and  $\gamma = \gamma_{\text{new}}$ . Otherwise, if membership (7) is false, another 1 is added:

$$\gamma_{\text{new}} = \frac{\gamma + \varphi(a) + 2}{2} \quad (9)$$

and  $\gamma = \gamma_{\text{new}}$ . Thus the new value of  $\gamma$  is always closer to the obsolete value of  $\gamma$  than to  $a$ . This is illustrated with Figure 2. An exception occurs when  $a$  happens to be integer and distance between the current value of  $\gamma$  and  $a$  is an even number. Then the new value of  $\gamma$  is exactly in the middle between the obsolete value of  $\gamma$  and  $a$ .



**Figure 2. An illustration of why the integer part of  $a$  is increased by either 1 (upper sketch) or 2 (lower sketch), while it gives the same  $\gamma_{\text{new}} = \gamma - 1$ ; in the uttermost case, when  $a$  goes to  $\gamma - 4$ ,  $\gamma_{\text{new}}$  is almost thrice closer to the obsolete value of  $\gamma$  than to  $a$**

If inequality (1) is false then it implies that the right end has been taken too small. The simplest correction is to drag it to the top right. So if

$$\frac{\gamma_{\max} + \varphi(a) + 1}{2} \in \mathbb{N} \quad (10)$$

then

$$\gamma = \frac{\gamma_{\max} + \varphi(a) + 1}{2} \quad (11)$$

Otherwise, if membership (10) is false, another 1 is added, similarly to that in (9):

$$\gamma = \frac{\gamma_{\max} + \varphi(a) + 2}{2} \quad (12)$$

An algorithm of such simple adjustment (hard adjustment, according with the said top-right-correction) uses the counter  $q^*$  for number of a-row-cases when  $a$  is approximately equal to  $\gamma$ , i. e. inequality (1) holds along with that either inequality (5) or inequality (6) is true (Figure 3). This counter is set back to zero if either inequality (1) is false or both inequalities (5) and (6) occur false. As soon as a row of  $Q^*$  cases in which  $a$  happened approximately to be equal to  $\gamma$  is scored, the algorithm states that  $\gamma$  is adjusted owing to that (the integer part of)  $a$  is too close to  $\gamma$  for the  $Q^*$ -th consecutive time. This is an approximation to statistical regularity [3, 4, 6, 7, 11]. It looks like a reason for early stop. And it becomes a statistically valid stopper for sufficiently great  $Q^*$ .

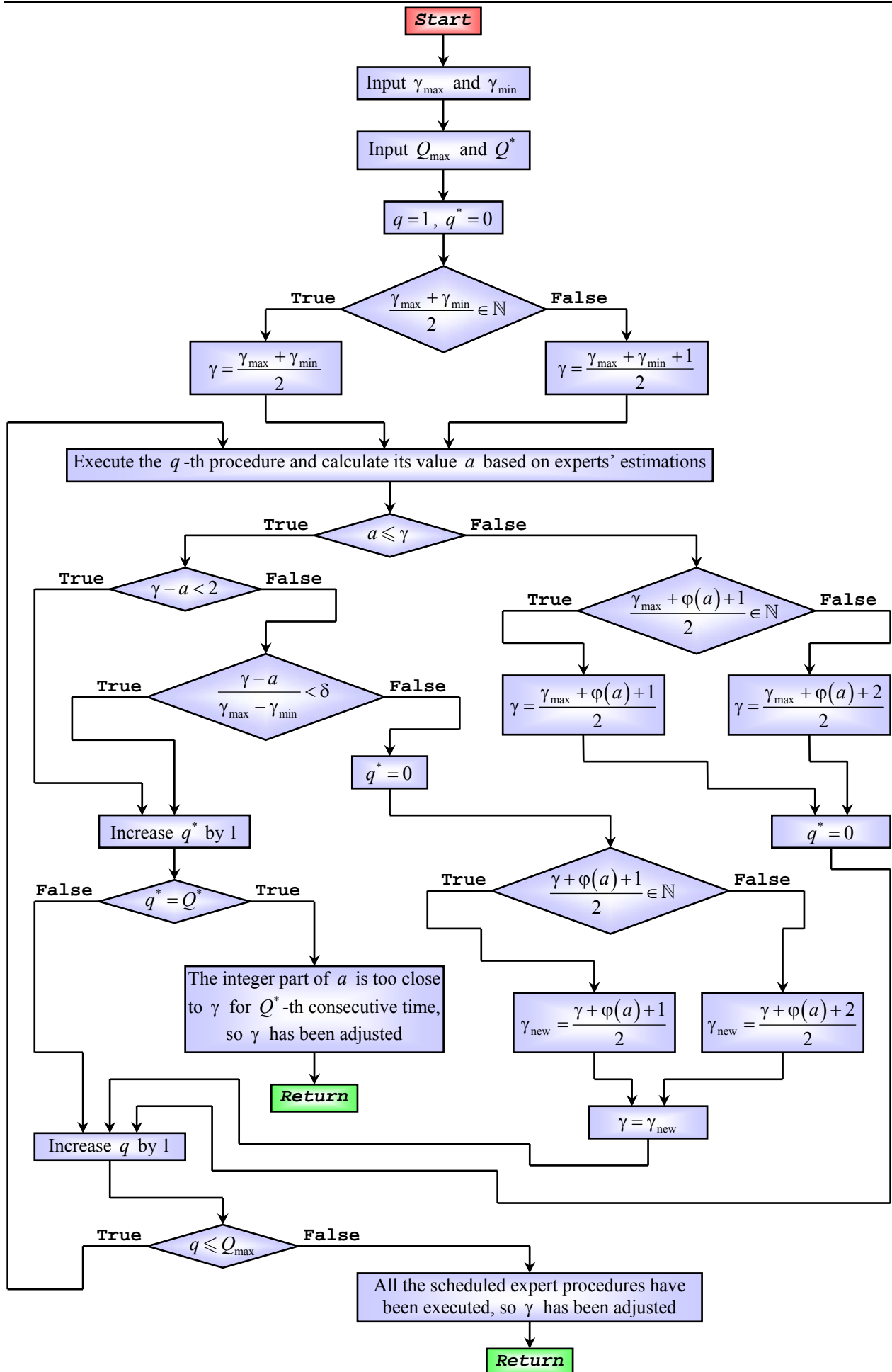


Figure 3. An algorithm of the positive integer parameter hard adjustment (tolerance is preliminarily assigned by default)

The second possibility of the algorithm's return is to pass through all the procedures. Value of  $\gamma$  is said to have been adjusted when all the schedule expert procedures are executed. Such a conclusion relies on that number  $Q_{\max}$  is assigned appropriately great. However,  $Q_{\max}$  may be assigned very small under the corresponding constrained conditions (short time, difficulties of expert procedure organization, low-confident experts, and so forth). Then, anyway, the algorithm's return should be accepted as it is, because of the conditions which cannot be amended straight off.

### Soft adjustment of the integer parameter

The algorithm in Figure 3 appears really rough when inequality turns false for  $a$  being much closer to  $\gamma_{\min}$ , and thus it is much farther from  $\gamma_{\max}$ . For avoiding such situations leading to paradoxical corrections, we can take the appropriate mean between  $a$  and some value of  $\gamma$  which was previously set to be the current parameter's value. This is going to be a modification of the hard adjustment algorithm.

Let enumerate all values of  $\gamma$  which were previously set, but now they are already obsolete. As it was mentioned above, the first obsolete value is  $\gamma_{\max}$ , i. e.

$$\gamma_{\text{obsolete}}^{(1)} = \gamma_{\max}.$$

At the start, the counter for obsolete values is set to 1.

For the modification of the hard adjustment algorithm in Figure 3, the algorithmic branch with the true inequality (1) remains almost the same. If inequalities (5) and (6) are false, then either (8) or (9) is calculated depending on membership (7). But before setting  $\gamma = \gamma_{\text{new}}$ , the obsolete value counter is increased by 1 and the next obsolete value is given:

$$\gamma_{\text{obsolete}}^{(k)} = \gamma.$$

If inequality (1) is false, then inequality

$$a \leq \gamma_{\text{obsolete}}^{(k)} \quad (13)$$

is checked next. The obsolete value counter is decreased by 1 while inequality (13) turns false. At this stage, the obsolete values, whose counter is decreased, are flushed. Decreasing is nonetheless limited owing to

$$a \leq \gamma_{\max} = \gamma_{\text{obsolete}}^{(1)}. \quad (14)$$

Inequality (14) is the worst case occurring after inequality (13) has turned false for a row of cases. This is why no explicit restriction on the obsolete value counter decrement is imposed. As soon as inequality (13) is true, closeness of value  $a$  to  $\gamma_{\text{obsolete}}^{(k)}$  is checked similarly to checks with inequalities (5) and (6). If

$$\gamma_{\text{obsolete}}^{(k)} - a < 2 \quad (15)$$

then  $\gamma = \gamma_{\text{obsolete}}^{(k)}$  by  $q^* = 1$  and the next expert procedure is executed, if only number of the already executed procedures is  $q < Q_{\max}$ . If inequality (15) is false then it is checked on whether  $a$  is approximately equal to  $\gamma_{\text{obsolete}}^{(k)}$  (or whether they are too close) using the same tolerance  $\delta$ . If

$$\frac{\gamma_{\text{obsolete}}^{(k)} - a}{\gamma_{\max} - \gamma_{\min}} < \delta \quad (16)$$

then, again,  $a$  is said to be approximately equal to  $\gamma_{\text{obsolete}}^{(k)}$  (these points are too close), and the current value of  $\gamma$  is set to  $\gamma = \gamma_{\text{obsolete}}^{(k)}$ . If inequality (16) is false, then, due to the right-side priority, membership

$$\frac{\gamma_{\text{obsolete}}^{(k)} + \varphi(a) + 1}{2} \in \mathbb{N} \quad (17)$$

is checked. If membership (17) holds then

$$\gamma = \frac{\gamma_{\text{obsolete}}^{(k)} + \varphi(a) + 1}{2}. \quad (18)$$

Otherwise, if membership (17) is false, another 1 is added, similarly to that in (9) and (12):

$$\gamma = \frac{\gamma_{\text{obsolete}}^{(k)} + \varphi(a) + 2}{2}. \quad (19)$$

After a new value of  $\gamma$  is calculated by either (18) or (19), the counter for number of a-row-cases, when  $a$  is approximately equal to  $\gamma$ , is set back to zero.

An algorithm representing the described adjustment of the positive integer parameter is softer (Figure 4)

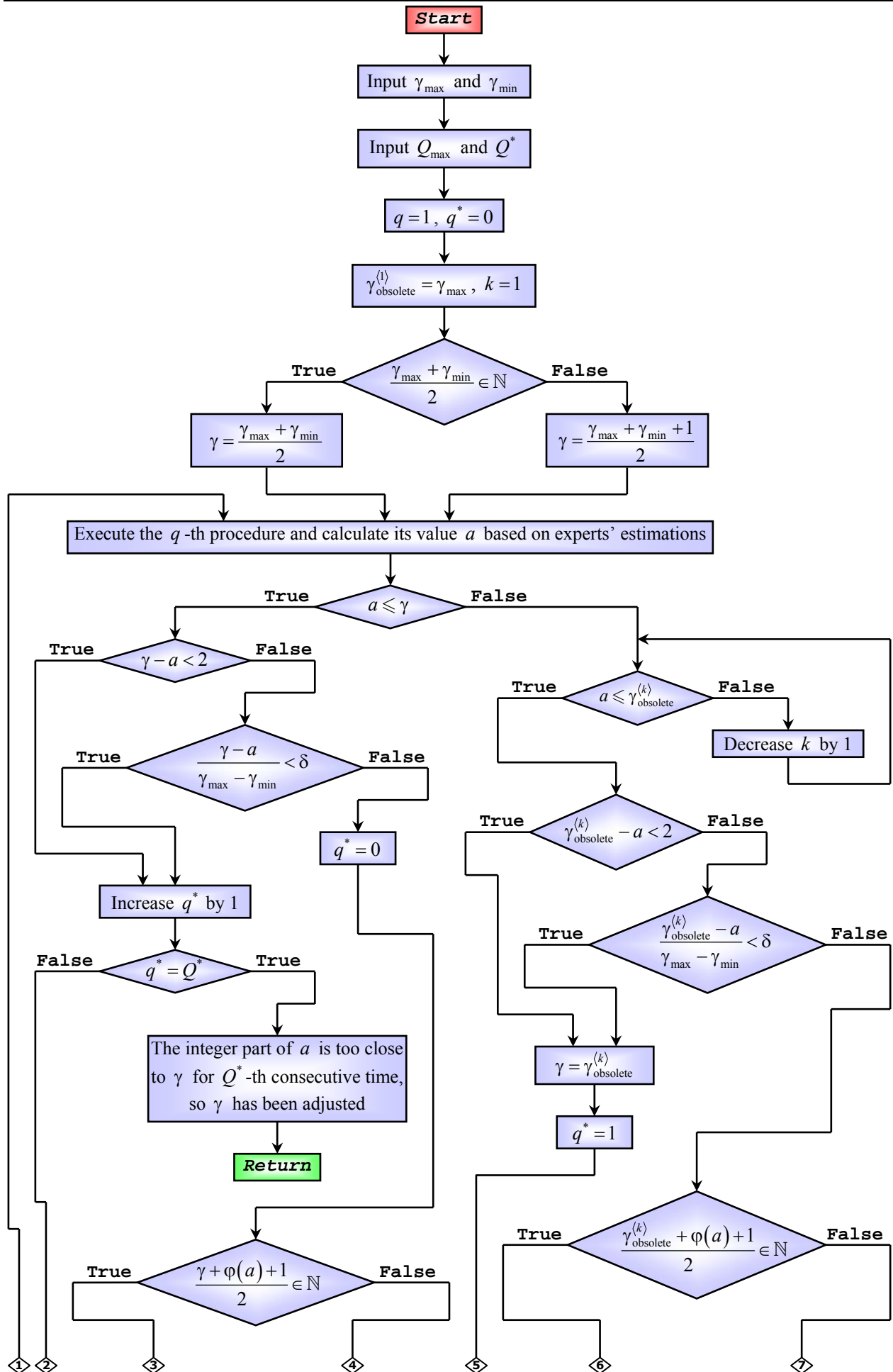


Figure 4. An algorithm representing soft adjustment of the positive integer parameter (beginning)

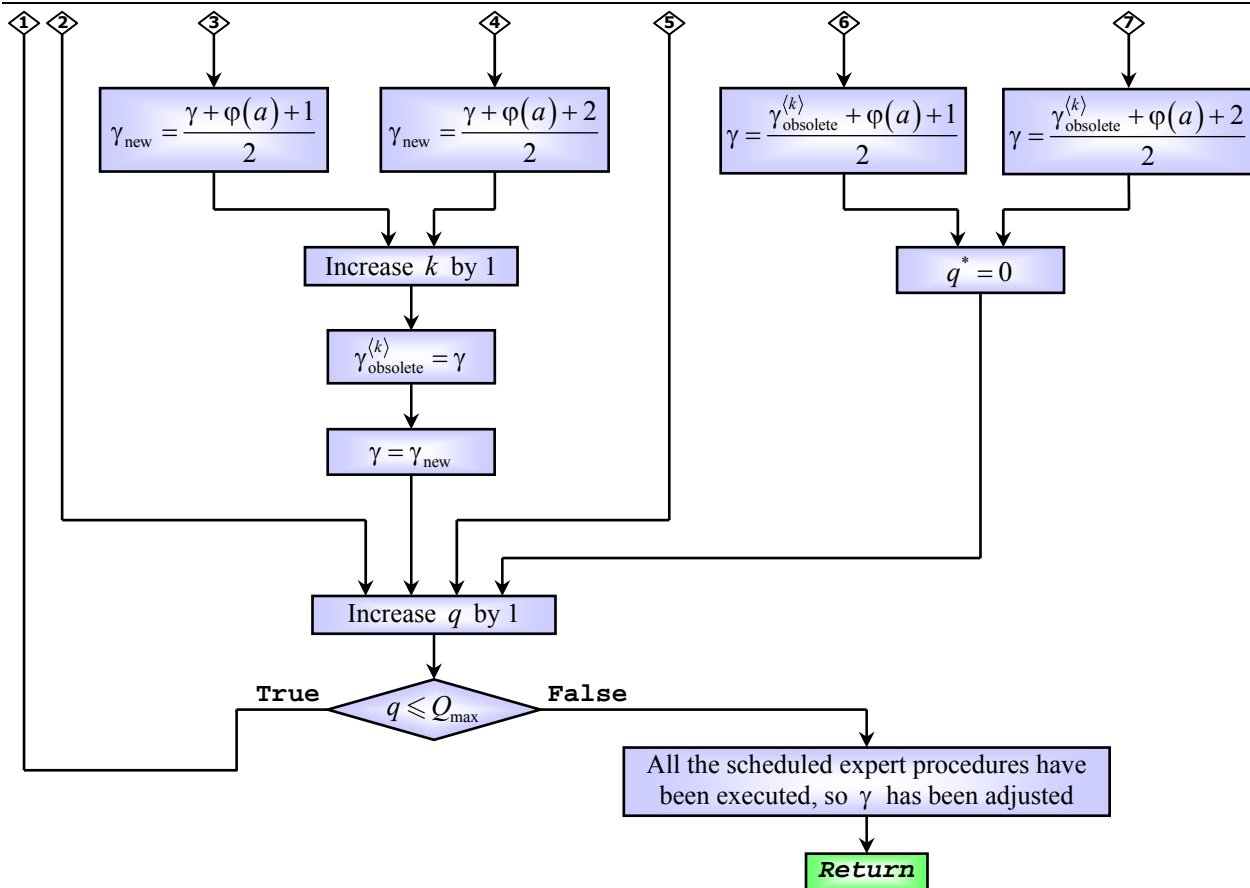


Figure 4. An algorithm representing soft adjustment of the positive integer parameter (completion)

when value  $a$  comes greater than the parameter, though earlier inequality (1) was true. The simplest correction dragging it to the top right happens only if inequality (13) turns false for any obsolete value of  $\gamma$  excepting the very first obsolete value  $\gamma_{\max}$  giving always-true inequality (14). The soft correction and the algorithmic branch with the false inequalities (5) and (6) adding the new obsolete value both constitute the modification.

### Application of the developed algorithms

The hard adjustment should precede the soft one. If some unknown integer parameter cannot be greater than an integer limit, then the developed algorithms fit. For integer hyperparameters of neural networks, an expert procedure is the training and testing on performance (error rate or accuracy). With the current value of  $\gamma$ , a group of neural networks is trained under various integers whose average is roughly equal to  $\gamma$ . Then they are tested, and the worst approximators are excluded, while the best ones are held, and average of their integers is assigned to  $a$ .

Such an application is usual for statistical approximators. Applications in other fields need specific interpretations, unless  $a$  is an average given straightforwardly by experts' estimations. For non-straightforward obtaining of  $a$ , the common feature of the interpretations is a criterion of the estimations, whatever field is. The criterion for neural networks is their performance.

The hard adjustment algorithm should be applied in fields of poor knowledge or with less proficient experts. On the contrary, the soft adjustment algorithm fits strong experience and highly proficient experts. Unfortunately, any strict criterion for choosing either the hard or soft adjustment is incomprehensible yet.

### Discussion

A substantial demerit of the hard adjustment is the-top-right-dragging back when inequality (1) turns false and  $a$  has been situated very close to  $\gamma_{\min}$ . So when  $a$  is approaching to  $\gamma_{\min}$ , the soft adjustment algorithm should supersede the hard one. Shifting back to the hard adjustment is possible.

Each algorithm may have its own tolerance  $\delta$ . But ratio between the hard and soft algorithmic tolerances is not definite. Tolerance in the hard algorithm may be as greater than that in the soft one, as well as lesser. Although formally, according with inequality (5) and the corresponding algorithmic branch logic,

$$\delta > 2 \cdot (\gamma_{\max} - \gamma_{\min})^{-1}$$

and  $\delta < 1$ , selection of tolerance is a separate task. To select it simpler, it may be taken as

$$\delta > 3 \cdot (\gamma_{\max} - \gamma_{\min})^{-1}$$

what is equivalent to the inequality  $\gamma - a < 3$  following the inequality (5) after it turned false.

Maximal number of expert procedures and number of a-row-cases, when  $a$  is approximately equal to  $\gamma$  by (1), are assigned beyond the algorithms. Their selection influences on tolerances. Nevertheless, for those numbers assigned sufficiently great, tolerance can be taken as desired. Moreover, the early stop condition becomes statistically valid. This saves time and expert resources.

### Conclusion and further work outline

Adjustment of a positive integer parameter is fulfilled by the conceptual condition (1) and subsequent mean dichotomy. Depending on the dichotomic style, hard and soft adjustments are discriminated and used supporting each other. Because of the average-like value of expert estimations is upper-limited to the parameter, dichotomy is fulfilled by the prioritized shift to the right (sketches in Figure 1 and Figure 2).

Algorithms in Figure 3 and Figure 4 are ready-to-run after initializing them with five input parameters. No strict requirements are imposed on those inputs. The boundaries can be assigned artificially exaggerated, letting the interval be looser. The looser interval expects greater number of expert procedures. This is especially suitable for fitting statistical approximators, where number of procedures implying training and testing may be as great as needed.

Further work concerns the inversion of how the parameter limits a result of expert estimations. Then inequality (1) will be reversed. The developed algorithms are the strong basis for solving the reverse adjusting problem. However, a case with limiting from both ends simultaneously is not going to be trivial.

### References

1. Reilly E. D. Milestones in Computer Science and Information Technology / E. D. Reilly. — Westport, CT : Greenwood Publishing Group, 2003. — 392 p.
2. Tucker A. B. Computer Science Handbook (2nd ed.) / A. B. Tucker. — Boca Raton, FL : Chapman and Hall/CRC, 2004. — 2624 p.
3. Blalock H. M. Social Statistics / H. M. Blalock. — New York : McGraw-Hill, 1979. — 592 p.
4. Manly B. F. J. Statistics for Environmental Science and Management / B. F. J. Manly. — Boca Raton, FL : Chapman & Hall/CRC, 2008. — 295 p.
5. LeCun Y. Gradient-based learning applied to document recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner // Proceedings of the IEEE. — 1998. — Vol. 86, Iss. 11. — P. 2278 — 2324.
6. Krizhevsky A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G. E. Hinton // Advances in Neural Information Processing Systems. — 2012. — Vol. 1. — P. 1097 — 1105.
7. Ciresan D. Flexible, high performance convolutional neural networks for image classification / D. Ciresan, U. Meier, J. Masci, L. M. Gambardella, J. Schmidhuber // Proceedings of the Twenty-Second international joint conference on Artificial Intelligence. — 2011. — P. 1237 — 1242.
8. Guo Q. Convolutional feature learning and Hybrid CNN-HMM for scene number recognition / Q. Guo, F. Wang, J. Lei, D. Tu, G. Li // Neurocomputing. — 2016. — Vol. 184. — P. 78 — 90.
9. Romanuke V. V. Setting the hidden layer neuron number in feedforward neural network for an image recognition problem under Gaussian noise of distortion / V. V. Romanuke // Computer and Information Science. — 2013. — Vol. 6, N. 2. — P. 38 — 54.
10. Gratie C. Complete characterization for the fit-preserving data refinement of mass-action reaction networks / C. Gratie, I. Petre // Theoretical Computer Science. — 2016. — Vol. 641. — P. 11 — 24.
11. DiCiccio T. J. Quantifying nuisance parameter effects via decompositions of asymptotic refinements for likelihood-based statistics / T. J. DiCiccio, T. A. Kuffner, G. A. Young // Journal of Statistical Planning and Inference. — 2015. — Vol. 165. — P. 1 — 12.
12. Xia S. The application of orthogonal test method in the parameters optimization of PEMFC under steady working condition / S. Xia, R. Lin, X. Cui, J. Shan // International Journal of Hydrogen Energy. — 2016. — Vol. 41, Iss. 26. — P. 11380 — 11390.
13. Marinoni M. Identification of groundwater flow parameters using reciprocal data from hydraulic interference tests / M. Marinoni, F. Delay, P. Ackerer, M. Riva, A. Guadagnini // Journal of Hydrology. — 2016. — Vol. 539. — P. 88 — 101.
14. Liljenherte J. Hyperelastic strain measurements and constitutive parameters identification of 3D printed soft polymers by image processing / J. Liljenherte, P. Upadhyaya, S. Kumar // Additive Manufacturing. — 2016. — Vol. 11. — P. 40 — 48.
15. Nunes L. C. S. Parameters identification of polymer concrete using a fracture mechanics test method and full-field measurements / L. C. S. Nunes, J. M. L. Reis, H. S. da Costa Mattos // Engineering Fracture Mechanics. — 2011. — Vol. 78, Iss. 17. — P. 2957 — 2965.
16. Zhou Y. Bayesian network approach to multinomial parameter learning using data and expert judgments / Y. Zhou, N. Fenton, M. Neil // International Journal of Approximate Reasoning. — 2014. — Vol. 55, Iss. 5. — P. 1252 — 1268.
17. Fu C. A group evidential reasoning approach based on expert reliability / C. Fu, J.-B. Yang, S.-L. Yang // European Journal of Operational Research. — 2015. — Vol. 246, Iss. 3. — P. 886 — 893.