Interval Computations
No 4(6), 1992

NE-FACTORS AND NATURAL PRAGMATICS: WHAT DO THE INTERVALS REPRESENT

Alexander S.Narin'yani

The interval notion can be used to represent several closely related but different features of the real world parameters, such as imprecision, subdefiniteness, uncertainness and fuzziness. In this paper, we discuss informally the difference in pragmatics of intervals representing these four aspects of quantitative knowledge. This difference is quite important when a formal computational model should adequately reflect real task with its natural inheritance of the imprecision and other mentioned factors. A well developed technology has been elaborated to deal successfully with subdefinite computational models.

НЕ-ФАКТОРЫ И ЕСТЕСТВЕННЫЙ ПРАГМАТИЗМ: ЧТО ПРЕДСТАВЛЯЮТ ИНТЕРВАЛЫ

А.С.Нариньяни

Понятие интервала может быть использовано для представления близких, но различных свойств параметров естественного мира, таких, как неточность, недоопределенность, нечеткость, неочевидность. В настоящей работе мы неформально обсуждаем различие в прагматике интервалов, представляющих эти четыре аспекта количественных знаний. Это различие является очень важным в случае, когда формальная вычислительная модель должна адекватно отражать реальную задачу с ее естественным наследованием неточности и других упомянутых факторов. Предлагается хорошо разработанная технология для успешной работы с недоопределенными вычислительными моделями.

1. In his paper [1], the author had made an attempt discuss comparatively several notions which are widely used in technical applications and knowledge engineering but rather poorly investigated or ignored at all in

the traditional mathematics. In this paper, they were called NE-factors from the Cyrillic HE, which is Russian equivalent for no-, non-, in-, un- and other negative prefixes, because both lexically and semantically, those real factors are opposite to the basic artificial features of the traditional mathematics such as definiteness, precision, completeness, correctness, etc.

Here, we shall discuss four NE-factors which can be represented using an interval data type. They are imprecision, subdefiniteness, uncertainness and fuzziness. The latter notion introduced by L.Zadeh in paper [2] has stabilized solidly during the previous almost three decades and born its own field of researches, the fuzzy mathematics which is having now all attributes of a serious science: a vast bibliography, specialized magazines, many annual conferences and even specific computers, based on the fuzzy logics.

This success would be very aspirating if not one small unpleasant "detail": every formal notion, even a fuzzy one, must not have a fuzzy pragmatics. It means that there should be defined most precisely what concrete features of real objects the formal apparatus represents, because these features do define the choice of operations and relations over the formal objects as well as the interpretation of given objects. In the opposite case, this choice becomes a rather random one: pretending to reflect the reality, the system turns to be a pseudo-model and leads to the results whose adequacy is impossible to check. In a great degree, this can be applied to the fuzzy apparatus, but a more extensive discussion of the subject is out of the framework of the present paper.

2. The same lack of definiteness of pragmatics should also be referred to the intervals. Really, the interval can represent the following:

(a) An *imprecise* value, i.e. such an estimation which can be given/acquired only with some particular precision, defined by the nature of a given parameter. It's obvious that all practical estimations are imprecise and that even an estimation of the precision itself is also imprecise. Any attempt to specify the imprecise value under higher precision has no physical sense. For example, an estimation of some river depth with the precision up to a centimeter. So we can say that the interval for an imprecise value is equal to the value it represents.

(b) A *subdefinite* value which is an estimation of a parameter having

a higher precision in principal but in the current situation being represented by a given wider interval due to rough measurements or/and incomplete data. In such a case the interval includes the real value and when additional information becomes available, it can constrict, reflecting the parameter value with an increasingly higher precision.

(c) An *uncertain* value, being by its nature more precise than an interval representing it (i.e. being a subdefinite value), cannot be estimated more definitely due to some objective causes. As a rule, such an interval value is being related with some distribution of a measure of probability, possibility, plausibility, etc., which allows taking decisions under uncertain data. This distribution can change when additional facts are being evaluated. In particular, an uncertain interval can become more narrow, if the corresponding measure on its ends becomes equal to zero.

(d) A *fuzzy* value which represents a linguistic variable, (also introduced by L.Zadeh), i.e. such notions as green, quickly, wide, often, etc. In the fuzzy mathematics, this variable is described with a fixed, i.e. static, interval over the corresponding scale of quantitative measurements (a spectrum, a scale of velocities, etc.). The fuzzy interval is associated with a function , also a static one, which defines for each value of the interval to what degree (expressed in the range [0,1]) this value belongs to the given notion.

It seems quite obvious that the described data type which is composed of a static interval and a static function cannot reflects adequately the pragmatics of concrete linguistic variables. To our opinion, it is more natural to consider these variables possessing to some degree the features of all three types of parameters, described in (a)–(c). For example, in the expression "not far from Moscow" the value "not_far" reflects a particular distance from Moscow of some concrete real object. This distance is imprecise by its own nature but, in a given context, it is estimated even more roughly. Value "not_far" covers this real estimation the author means and, consequently, it can become more definite, i.e. constrict its interval. Both the author and the listener (if he doesn't know the precise distance) have their own plausibility functions which estimate also imprecisely a possible position of the real distance value within the interval corresponding to the linguistic variable "not_far".

It seems obvious enough that with development of the knowledge representation apparatus new types of variables will be created which will reflect adequately several interrelated NE-factors. For example, to model the linguistic variable, described above. It can be represented with a pair of (i) a subdefinite value which can constrict to the degree limited by the imprecision of a given parameter and (ii) a function which is defined on this subdefinite interval and estimating plausibility of each (imprecise) value within the interval.

3. The arguments above can be considered as being abstract enough and remote from the real practice of the interval computations. The pragmatics of the computations seems perfectly transparent - you define intervals of the input parameters and receive at the output the intervals of the results.

It could appear that for these results that it is immaterial, which particular NE-factors the parameters of the computations represent: there is important only that these parameters are known with some concrete precision and the latter defines the precision of the output values.

But in the reality, the difference is rather great. As a rule, for every task in addition to the algorithm of computation, there are complimentary restrictions which bound the input and output parameters of the task and, possibly, a procedure defining the inverse function. In such a case intervals of the same variable which are obtained by different ways can be not identical but only intersecting. If this concerns a subdefinite parameter, then, under certain conditions, its value can be taken equal to the meet of the corresponding intervals. But in the case of imprecise variable, this situation means that the computations are not correct.

4. These considerations become especially actual when going from interval algorithms to interval models. In the models, there is no separation of the parameters into input and output ones. All of them are just a set of variables bounded with a sum of constraints, i.e. equations, inequalities and logic relations. For the interval models in which some parameters are subdefinite variables, the computational process can be organized in such a mode [3], that all these variables interact through the model and make each other increasingly more definite, constricting the space of solution up to intervals which satisfy all constraints of the model.

In this process, it's very important of course which particular NE-

factors the model's parameters represent because namely this defines to what limit the corresponding intervals can constrict. For example, the relations

$$x * (y - 1) = 1; \quad 2 * x + y \leq 4; \quad x > 0;$$

define the space $x = [0.49999;\ 1.00002]$, $y = [1.99996;\ 3.00004]$, if x and y are subdefinite variables, and the space $x = [0.45778;\ 1.09224]$, $y = [1.86555;\ 3.13447]$, if y is an imprecise parameter with its "width" equal to 0.1.

5. The approach presented above isn't a purely theoretical one. The joint team of the Russian Research Institute of Artificial Intelligence and scientific firm "Intelligent Technology" have developed and implemented the apparatus of active data types [4] which allows constructing types of variables representing different NE-factors in the models, based on the data-driven computational process. In particular, these data types can be constructed not only on the basis of the simple interval: for example, it's possible to build such ones whose current subdefinite value is represented with a set of intervals or even with a set of individual values (for integers).

This apparatus and the library of subdefinite data types permit the research group to create a spectrum of next generation software systems and technologies which implement qualitatively new possibilities in a wide range of applications [5].

References

1. Narin'yani, A.S. *Sub-definiteness in knowledge representation and processing.* Technical cybernetics **5** (1986), pp. 3–28 (in Russian).
2. Zadeh, L.A. *Fuzzy sets.* Inf. & Control **8** (1965).
3. Narin'yani, A.S., Telerman, V.V. and Dmitriev, V.E. *Virtual data-flow machine as vehicle of inference/computations in knowledge bases.* In: Jorrand, Ph. and Sgurev, V. (eds) "Artificial intelligence II. Methodology, systems, application", North-Holland, 1987, pp. 149–155.
4. Telerman, V.V. *Active data types.* Preprint 792, Computer Center, Siberian Branch of the USSR Academy od Sciencres, Novosibirsk (1988) (in Russian).
5. Narin'yani, A.S. *Intelligent software technology for the new decade.* Communications of the ACM **34** (6) (1991), pp. 60–67.