Chemical Kinetics in Situations Intermediate Between Usual and High Concentrations: Fuzzy-Motivated Derivation of the Formulas

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

Laécio Carvalho Barros
Instituto de Matemática, Estatística,
e Computação Científica (IMECC
Universidade Estadual de Campinas (UNICAMP)
Campinas, SP, Brasil C.P. 6065
laeciocb@ime.unicamp.br

Abstract—In the traditional chemical kinetics, the rate of each reaction

$$A + \ldots + B \to \ldots$$

is proportional to the product $c_A \cdot \ldots \cdot c_B$ of the concentrations of all the input substances A, ..., B. For high concentrations $c_A, \ldots, c_B$, the reaction rate is known to be proportional to the minimum $\min(c_A, \ldots, c_B)$. In this paper, we use fuzzy-related ideas to derive the formula of the reaction rate for situations intermediate between usual and high concentrations.

I. Chemical Kinetics in Situations Intermediate Between Usual and High Concentrations: Formulation of the Problem

Chemical kinetics: usual formulas. Chemical kinetics describes the rate of chemical reactions. For usual concentrations, the rate of a reaction between two substances A and B is proportional to the product $c_A \cdot c_B$ of their concentrations; see, e.g., [3], [7]. Similarly, if we have a reaction

$$A + \ldots + B \to \ldots$$
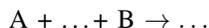
with three or more substances, the rate of this reaction is proportional to the products of the concentrations of all these substances $c_A \cdot \ldots \cdot c_B$.

How formulas of chemical kinetics are usually derived. Let us start the explanation of how the *general* formulas of chemical kinetics are derived by first considering the case of *two* substances A and B.

Molecules of both substances are randomly distributed in space. So, for each molecule of the substance A, the probability that it meets a molecule of the substance B is proportional to the concentration c_B. If the molecules meet, then (with a certain probability) they get into a reaction. Thus, the expected number of reactions involving a given molecule of the substance A is also proportional to c_B. The total number of A-molecules in a given volume is proportional to c_A; thus, the total number of reactions per unit time is proportional to $c_A \cdot c_B$.

Similarly, for the case of three or more substances, we can conclude that the reaction rate is indeed proportional to the product $c_A \cdot \ldots \cdot c_B$.

Case of high concentrations. When the concentrations are very high, there is no need for the molecules to randomly bump into each other; these molecules are everywhere. So, as soon as we have molecules of all needed type, the reaction starts. In other words, in this case, the reaction rate is proportional to the concentration of the corresponding tuples – i.e., to the minimum $\min(c_A, \ldots, c_B)$ of all the input concentrations $c_A, \ldots, c_B$.

Example. The formula $\min(c_A, \ldots, c_B)$ can be easily illustrated on the example of a relation which is non-chemical reaction but which is described by the same chemical kinetic-type equations: the relation between predators and prey.

When we have usual (small) concentrations of wolves W and rabbits R in a forest, the probability for a wolf to find a rabbit is proportional to the concentration c_R of rabbits, so the overall amount of rabbits eaten by wolves is proportional to the product $c_W \cdot c_R$.

On the other hand, for high concentrations, e.g., if we throw a bunch of rabbits into a zoo cage filled with hungry wolves, there is no need to look for a prey, each wolf will start eating its rabbit – as long as there are sufficiently many rabbits to feed all the wolves. So:

- When $c_R \geq c_W$, the number of eaten rabbits will be proportional to the number of wolves, i.e., to c_W.
- In situations when there are not enough rabbits (i.e., when $c_R < c_W$), the number of eaten rabbits is proportional to the number of rabbits, i.e., to c_R.

In both cases, the reaction rate is proportional to $\min(c_R, c_W)$.

Empirical evidence for high-concentration reaction rate. The high-concentration reaction rate indeed turned out to be very useful to describe biochemical processes; see, e.g., [2], [8].

Interesting observation: simulations of high-concentration reactions lead to efficient algorithms. It is known that in many cases, difficult-to-solve computational problems can be reduced to problems of chemical kinetics. In such situations of *chemical computing*, we can efficiently solve the original

computational problems by either actually performing the corresponding chemical reactions, or by performing a computer simulation of these reactions; see, e.g., [1].

To make the simulations as fast as possible, it is desirable to simulate reactions which are as fast as possible. The reaction rate increases with the concentrations of the reagents. Thus, to speed up simulations, we should simulate high-concentration reactions. This simulation indeed speeds up the corresponding computations; see, e.g., [5], [6].

Main problem. While we know the formulas for the usual and for the high concentrations, it is not clear how to compute the reaction rate for concentrations between usual and high.

What is known. Both formulas $r = c_A \cdot c_B$ and $r = \min(c_A, c_B)$ are particular cases of *t-norms* ("and"-operations in fuzzy logic; see, e.g., [4], [9], [11]). This is not a coincidence: there is no reaction if one of the substances is missing, so $c_A = 0$ or c_B imply that $r = 0$ – which is exactly the property of a t-norm. Fuzzy t-norms have indeed been effectively used to describe chemical reactions [2], [8].

Remaining problem. The problem is that there are many possible "and"-operations, and it is not clear which one we should select.

What we do in this paper. In this paper, we use the analysis of the corresponding chemical processes to derive the formulas that adequately describe the reaction rate in intermediate situations – and thus, to appropriately select the corresponding "and"-operation.

II. Chemical Kinetics in Situations Intermediate Between Usual and High Concentrations: Analysis of the Problem, Resulting Formulas, and Discussion

Towards formulating the problem in precise terms. Let us start with the case of two substances A and B. As we have mentioned earlier, the two molecules get into a reaction only when they are close enough. When these molecules are close enough, then, within the corresponding small region, the reaction rate is proportional to the minimum $\min(c_A, c_B)$ of their concentrations.

When concentrations are small, then, within each region, we have either zero or one molecule; the probability to have two molecules is very small (proportional to the square of these concentrations) and can, therefore, be safely ignored. In this case, for each region, the reaction occurs if we have both an A-molecule and a B-molecule. The probability to have an A-molecule is proportional to c_A; the probability to have a B-molecule is proportional to c_B. Since the distributions for A and B are independent, the probability to have both A- and B-molecules in a region is equal to the product of these probabilities and is, thus, proportional to the product of the concentrations $c_A \cdot c_B$.

When the concentrations are high, then each region has molecules of both types. The average number of A-molecules in a region is proportional to c_A, i.e., has the form $k \cdot c_A$

for some proportionality coefficient k. Similarly, the average number of B-molecules in a region is equal to $k \cdot c_B$. So the average reaction rate is proportional to $\min(k \cdot c_A, k \cdot c_B) = k \cdot \min(c_A, c_B)$, i.e., is proportional to $\min(c_A, c_B)$.

This analysis leads us to the following reformulation of our problem.

Resulting formulation of the problem in precise terms. Within a unit volume, we have a certain number r of " small regions", i.e., regions such that only molecules within the same region can interact with each other.

We have a total of $N_A = N \cdot c_A$ molecules of the substance A, and we have a total of $N_B = N \cdot c_B$ molecules of the substance B. Each of these molecules is randomly distributed among the regions, i.e., it can be located in any of the r regions with equal probability. Distributions of different molecules are independent from each other. Within each region, the reaction rate is proportional to the minimum $\min(n_A, n_B)$ of the numbers n_A and n_B of A- and B-molecules in this region. The overall reaction rate can be computed as the *average* over all the regions – i.e., in other words, as the mathematical expectation of this minimum.

Analysis of the problem. Based on the above description, the number of A-molecules in a region follows the Poisson distribution (see, e.g., [10]), according to which, for each value k, the probability to have exactly $n_A = k$ A-molecules is equal to

$$\text{Prob}(n_A = k) = \exp(-\lambda_A) \cdot \frac{\lambda_A^k}{k!}. \quad (1)$$

The mean value of the Poisson random variable is λ_A; on the other hand, we have $N \cdot c_A$ A-molecules in r cells, so the average number of A-molecules in a cell is equal to the ratio $\dfrac{N \cdot c_A}{r}$, so

$$\lambda_A = \frac{N \cdot c_A}{r}. \quad (2)$$

In other words, $\lambda_A = c \cdot c_A$, where we denoted $c \stackrel{\text{def}}{=} \dfrac{N}{r}$. Similarly, for the number n_B of B-molecules in the region, we have a probability distribution

$$\text{Prob}(n_B = k) = \exp(-\lambda_B) \cdot \frac{\lambda_B^k}{k!}, \quad (3)$$

with

$$\lambda_B = \frac{N \cdot c_B}{r} = c \cdot c_B. \quad (4)$$

The desired distribution for $n = \min(n_A, n_B)$ can be obtained from the fact that

$$n \geq k \Leftrightarrow (n_A \geq k \,\&\, n_B \geq k).$$

Since A- and B-molecules are independently distributed, the A-related value n_A and the B-related value n_B are also independent. Therefore,

$$\text{Prob}(n \geq k) = \text{Prob}(n_A \geq k) \cdot \text{Prob}(n_B \geq k). \quad (5)$$

Based on (1) and (3), we conclude that

$$\text{Prob}(n_A \geq k) = \sum_{\ell=k}^{\infty} \text{Prob}(n_A = \ell) =$$

$$\exp(-\lambda_A) \cdot \sum_{\ell=k}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} \qquad (6)$$

and

$$\text{Prob}(n_B \geq k) = \sum_{\ell=k}^{\infty} \text{Prob}(n_B = \ell) =$$

$$\exp(-\lambda_B) \cdot \sum_{\ell=k}^{\infty} \frac{\lambda_B^{\ell}}{\ell!}. \qquad (7)$$

So, we conclude that

$$\text{Prob}(n \geq k) =$$

$$\exp(-(\lambda_A + \lambda_B)) \cdot \left(\sum_{\ell=k}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} \right) \cdot \left(\sum_{\ell=k}^{\infty} \frac{\lambda_B^{\ell}}{\ell!} \right). \qquad (8)$$

The expected value E can be now computed as

$$E = \sum_{k=0}^{\infty} k \cdot \text{Prob}(n = k) =$$

$$\sum_{k=0}^{\infty} k \cdot (\text{Prob}(n \geq k) - \text{Prob}(n \geq k+1)) =$$

$$0 \cdot (\text{Prob}(n \geq 0) - \text{Prob}(n \geq 1)) +$$

$$1 \cdot (\text{Prob}(n \geq 1) - \text{Prob}(n \geq 2)) +$$

$$2 \cdot (\text{Prob}(n \geq 2) - \text{Prob}(n \geq 3)) + \ldots =$$

$$\text{Prob}(n \geq 1) \cdot (1 - 0) + \text{Prob}(n \geq 2) \cdot (2 - 1) +$$

$$\ldots =$$

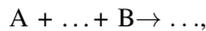$$\sum_{k=1}^{\infty} \text{Prob}(n \geq k). \qquad (9)$$

Substituting the expression (8) into this formula, we arrive at the following expression.

Resulting formula for the reaction rate. The reaction rate is proportional to

$$E \stackrel{\text{def}}{=} \exp(-(\lambda_A + \lambda_B)) \cdot \sum_{k=1}^{\infty} \left(\sum_{\ell=k}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} \right) \cdot \left(\sum_{\ell=k}^{\infty} \frac{\lambda_B^{\ell}}{\ell!} \right), \quad (10)$$

where $\lambda_A = c \cdot c_A$ and $\lambda_B = c \cdot c_B$ for some constant c.

For a reaction between three or more substances

$$A + \ldots + B \rightarrow \ldots,$$

we similarly get a formula

$$E \stackrel{\text{def}}{=}$$

$$\exp(-(\lambda_A + \ldots \lambda_B)) \cdot \sum_{k=1}^{\infty} \left(\sum_{\ell=k}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} \right) \cdot \ldots \cdot \left(\sum_{\ell=k}^{\infty} \frac{\lambda_B^{\ell}}{\ell!} \right), \quad (10a)$$

where $\lambda_A = c \cdot c_A, \ldots$, and $\lambda_B = c \cdot c_B$ for some constant c.

Towards simplifying the above formula. Let us show that the above formula can be somewhat simplified by expressing it in terms of the upper incomplete Gamma-function. The upper incomplete Gamma function is often used to analyze the Poisson distribution. It is defined as

$$\Gamma(s, x) \stackrel{\text{def}}{=} \int_x^{\infty} t^{s-1} \cdot \exp(-t) \, dt. \qquad (11)$$

Its relation to the Poisson distribution comes from the fact that for integer values s, we have

$$\exp(-\lambda) \cdot \sum_{\ell=0}^{s-1} \frac{\lambda^{\ell}}{\ell!} = \frac{\Gamma(s, \lambda)}{(s-1)!}. \qquad (12)$$

Since $\exp(\lambda) = \sum_{\ell=0}^{\infty} \frac{\lambda^{\ell}}{\ell!}$, we have

$$\exp(-\lambda) \cdot \sum_{\ell=0}^{\infty} \frac{\lambda^{\ell}}{\ell!} = 1. \qquad (13)$$

Subtracting (12) from (13), we conclude that

$$\exp(-\lambda) \cdot \sum_{\ell=s}^{\infty} \frac{\lambda^{\ell}}{\ell!} = 1 - \frac{\Gamma(s, \lambda)}{(s-1)!}. \qquad (14)$$

In particular, for $\lambda = \lambda_A$ and for $\lambda = \lambda_B$, we get the following formulas:

$$\exp(-\lambda_A) \cdot \sum_{\ell=s}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} = 1 - \frac{\Gamma(s, \lambda_A)}{(s-1)!}; \qquad (15)$$

$$\exp(-\lambda_B) \cdot \sum_{\ell=s}^{\infty} \frac{\lambda_B^{\ell}}{\ell!} = 1 - \frac{\Gamma(s, \lambda_B)}{(s-1)!}. \qquad (16)$$

Substituting these expressions instead of the sums into the formula (10), we arrive at the following expression.

Simplified version of the rate formula. The reaction rate is proportional to

$$E = \sum_{k=1}^{\infty} \left(1 - \frac{\Gamma(k, \lambda_A)}{(k-1)!} \right) \cdot \left(1 - \frac{\Gamma(k, \lambda_B)}{(k-1)!} \right). \qquad (17)$$

For the reaction between three or more substances, a similar formula takes the form

$$E = \sum_{k=1}^{\infty} \left(1 - \frac{\Gamma(k, \lambda_A)}{(k-1)!} \right) \cdot \ldots \cdot \left(1 - \frac{\Gamma(k, \lambda_B)}{(k-1)!} \right). \qquad (17a)$$

Analysis of the above formula. Let us show that in both limit cases – when concentrations are small and when concentrations are large – the formula (10) (and thus, the equivalent formula (17)) leads to the known expressions for the reaction rate.

Indeed, when λ_A and λ_B are small, then $\exp(-(\lambda_A + \lambda_B))$ is approximately equal to 1. Also, terms proportional to λ_A^2 and to higher powers of λ_A are much smaller than the term proportional to λ_A and can, therefore, be ignored. So, in this case, we have $\sum_{\ell=1}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} \approx \lambda_A$ and $\sum_{\ell=k}^{\infty} \frac{\lambda_A^{\ell}}{\ell!} \approx 0$ for $k > 1$.

Similarly, we have $\sum_{\ell=1}^{\infty} \dfrac{\lambda_B^\ell}{\ell!} \approx \lambda_B$ and $\sum_{\ell=k}^{\infty} \dfrac{\lambda_A^\ell}{\ell!} \approx 0$ for $k > 1$. Thus, the formula (10) takes the form $E = \lambda_A \cdot \lambda_B$. Since $\lambda_A = c \cdot c_A$ and $\lambda_B = c \cdot c_B$, this means that in this case, the reaction rate is indeed proportional to $c_A \cdot c_B$.

The estimate for the case when λ_A and λ_B are small was based on the fact that in this case, the terms $\dfrac{\lambda_A^\ell}{\ell!}$ drastically decrease with ℓ, so we only need to take the into account the largest term – which corresponds to the smallest possible value $\ell = 1$. When λ_A and λ_B are large, the dependence on ℓ is no longer monotonic. The largest value of this term can be estimated if we approximate $\ell!$ by the usual Stirling approximation $\ell! \approx \left(\dfrac{\ell}{e}\right)^\ell$, reducing each term $\dfrac{\lambda^\ell}{\ell!}$ to $\left(\dfrac{\lambda \cdot e}{\ell}\right)^\ell$. This term is the largest when its logarithm

$$L \overset{\text{def}}{=} \ell \cdot (\ln(\lambda) + 1 - \ln(\ell))$$

attains the largest possible value. Differentiating L with respect to ℓ and equating the resulting derivative to 0, we conclude that $\ell = \lambda$. For this value ℓ, the term $\left(\dfrac{\lambda \cdot e}{\ell}\right)^\ell$ turns into $\exp(\lambda)$. Since $\exp(\lambda)$ is equal to the whole sum $\sum_{\ell=0}^{\infty} \dfrac{\lambda^\ell}{\ell!}$, this means that all other terms in this sum are much smaller – and can thus be, in the first approximation, ignored.

In this first approximation, we can therefore assume that this term is equal to $\exp(\ell)$, while all other terms are 0s. Thus, the sum $\sum_{\ell=k}^{\infty} \dfrac{\lambda_A^\ell}{\ell!}$ is equal to 0 when $\ell > \lambda_A$ and to $\exp(\lambda_A)$ when $\ell \leq \lambda_A$. Similarly, the sum $\sum_{\ell=k}^{\infty} \dfrac{\lambda_B^\ell}{\ell!}$ is equal to 0 when $\ell > \lambda_B$ and to $\exp(\lambda_B)$ when $\ell \leq \lambda_B$.

So, in the sum (10), the only non-zero terms correspond to cases when $\ell \leq \lambda_A$ and $\ell \leq \lambda_B$, i.e., when $\ell \leq \min(\lambda_A, \lambda_B)$. Each of these $\min(\lambda_A, \lambda_B)$ non-zero terms is equal to

$$\exp(-(\lambda_A + \lambda_B)) \cdot \exp(\lambda_A) \cdot \exp(\lambda_B) = 1,$$

so their sum is indeed approximately equal to $\min(\lambda_A, \lambda_B)$.

Remaining open questions. Formulas similar to chemical kinetics equations are used in many different applications, e.g., in the dynamics of biological species or in the analysis of knowledge propagation. In all these cases, we can consider the product and min operations, and we can also consider intermediate cases.

The above derivation of the formulas for the intermediate "and"-operation uses the specifics of chemical kinetics. It would be interested to provide a similar analysis in other applications areas and see which "and"-operations are appropriate in these situations.

References

[1] A. I. Adamatzky, "Information-processing capabilities of chemical reaction-diffusion systems. 1. Belousov-Zhabotinsky media in hydrogel matrices and on solid supports", *Advanced Materials for Optics and Electronics*, 1997, Vol. 7, No. 5, pp. 263–272.

[2] F. S. Pedro and L. C. Barros. *The use of t-norms in mathematical models of epidemics*, Fuzzy Systems (FUZZ), 2013 IEEE International Conference on. IEEE, 2013.

[3] J. E. House, *Principles of Chemical Kinetics*, Academic Press, Burlington, Massachisetts, 2007.

[4] G. Klir and B. Yuan, "Fuzzy Sets and Fuzzy Logic", Prentice Hall, Upper Saddle River, New Jersey, 1995.

[5] V. Kreinovich and L. O. Fuentes, "Simulation of chemical kinetics – a promising approach to inference engines, In: J. Liebowitz (ed.), *Proceedings of the World Congress on Expert Systems, Orlando, Florida, 1991*, Pergamon Press, New York, 1991, Vol. 3, pp. 1510–1517.

[6] V. Kreinovich and O. Fuentes, "High-concentration chemical computing techniques for solving hard-to-solve problems, and their relation to numerical optimization, neural computing, reasoning under uncertainty, and freedom of choice", In: E. Katz (ed.), *Molecular and Supramolecular Information Processing: From Molecular Switches to Logical Systems*, Wiley-VCH, Wienheim, Germany, 2012, pp. 209–235.

[7] G. Marin and G. S. Yablonsky, *Kinetics of Chemical Reactions*, Wiley-VCH, Weinheim, Germany, 2011.

[8] E. Massad, N. Ortega, L. Barros, and C. Struchiner, *Fuzzy Logic in Action: Applications in Epidemiology and Beyond*, Springer Verlag, Berlin, Heidelberg, 2008.

[9] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[10] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

[11] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.