Archives of Control Sciences
Volume 12(XLVIII), 2002
No. 4, *pages 457–467*

Granular entropy and granulation process

ALEXANDER RYBALOV

People use granulation to represent original data as a set of entities that are better suited for managing resulting subtasks. Concepts that are introduced in this paper are based on two features of granulation: the relevance of all the points in a granule and relative number of points in every granule. Based on these features several concepts are introduced that include granular relevance, defined as sum of relevancies of all the points in the granule and granular entropy, that is similar to information entropy and reflects the dispersion of relevant points across granules. Using these concepts granulation process is represented as solution of optimization problem where objective function is granular entropy. To this end the theorem, that shows how the change in relevance of points in a granule affects granular entropy, was proved. The last two sections of the paper show how leverage over the granulation process can be achieved by using t-norm and uni-norm operators.

Key words: granulation, entropy, clustering, control, t-norm, uni-norm operators

1. Introduction

The result of granulation process is the representation of basic information as a set of entities that are better suited for managing resulting subtasks. As granules represent concepts that humans use for different subtasks, the granulation process often lacks formal mechanism. As a result, it often happens that resulting granules are either too big or too small. Here we face paradoxical problem: on one hand granulation process is knowledge-oriented, i.e. is dependent on human understanding, and on other hand is data-driven, i.e. we don't have complete control over this process. As a result proper leveraging mechanism is lacking. To built it we need to utilize two sides of granulation process: importance of different points for a given problem and size of resulting granules. To achieve this we identify granules with clusters, and use process of clustering as constrain on granulation process. The latter can be presented, therefore, as optimization problem. What we need is proper representation.

As granulation process is essentially an information process, the notion of entropy used in information theory should be extended to granulation as well. Thus we get a measure that can be used to control granulation process. What we need at this stage is

The Author is with Jerusalem College of Technology, 21 HaVaad Haleumi st., Jerusalem, Israel, tel: (02) 675-1148, e-mail: alxndr_r@yahoo.com.

Received 29.10.2001, revised 8.04.2002.

to determine how granular entropy is changing during the course of granulation process. The next step is to determine what operator can be used to leverage this process. What we need is a class of flexible and consistent operators that permit us to control size and relevance of resulting granules. These operators should satisfy certain properties, in particular to be commutative, associative, and monotonic. Additional requirements should let us to complete leverage over granulation process.

2. Granular entropy and granulation process

Solution of problems requires division of initial problem into subtasks that are more suitable to human understanding. Some points are more relevant for solution of original problem, while others don't attain much importance. If some points are indispensable for solution of the original problem, i.e. their removal influence the results in *every* model, we can say that these points are highly relevant, or their relevancy is equal to 1. If, on other hand, removal of points in almost all possible models don't have effect on the solution then we can say that these points are irrelevant, or their relevancy is close to 0. To formalize this we define that every point p has **relevancy** $R_p \in (0,1]$. Experts can decide how relevant is each point and assign relevancy values for every point or a group of points. For example, in a space with six points first, second and third points, that form compact subspace, have relevancies equal to 0.4, 0.3 and 0.2, whereas fourth, fifth and sixth points, that also form compact subspace, have relevancies equal to 0.7, 0.8 and 0.9. If points with similar relevancy form compact subspace then we assume that they are within the same granule. So, in the above example, first, second and third points are in one granule, and fourth, fifth and sixth points are in another granule. Sum of all relevancies in a granule is called **granular relevance** G_j. It indicates how important each particular granule is. In the above example granular relevancies are equal to 0.9 for the first granule, and 3.3 for the second granule, and, as a result, we can see that the second granule is much more important than the first. Sum of granular relevancies for all granules is called **total granular relevance**. In the example above it is equal to 4.2.

The formation of granules should depend on dispersion of points with different relevancies across granules and distribution of granular relevancies. To get a measure of dispersion of points across granules we introduce the notion of **granular entropy**.

Definition: *If there are m granules where each granule j having $S(j)$ points, and each point i has relevancy R_i, then* **granular entropy** *GE is defined as:*

$$GE = -\sum_{j=1}^{m} \frac{G_j}{G} \cdot \ln\left(\frac{G_j}{G}\right), \quad \text{where} \quad G_j = \sum_{i \in S(j)} R_i, \quad G = \sum_{j=1}^{m} G_j. \tag{1}$$

G_j is granular relevance, and G is total granular relevance. The value of granular entropy in the above example is equal to 0.5196.

This measure is similar to the measure of information entropy and gives us a tool to leverage the granulation process. Essentially it measures the dispersion of the most relevant (and important) points across granules: the larger dispersion is the higher granular entropy is. As points in granules form compact subspaces, we can have a choice in granulation process. If we aim at maximizing granular entropy (with number of granules fixed) we will get granules that have approximately equal number of highly relevant points. This can significantly enhance our understanding of initial problem's space and help to divide it into appropriate subtasks. If, on other hand we want to minimize granular entropy, we will get one granule that have all points that are relevant, and the rest dispersed among other granules. Some applications do require such type of granulation.

But there are dangers in the first approach. Maximization of granular entropy can lead to the situation where small number of points, that have high relevancy are spread over number of granules, whereas all other points are concentrated within one granule. In this case granular entropy will be high, but at the expense of conceptual understanding of the problem. Insight into essence of the problem is required as well, and for this we need carefully calibrate granulation process.

To achieve this at first we identify granules with clusters (i.e. they have the same points). Next we introduce the notion of **cluster entropy**.

Definition: *Suppose we have m clusters with n_i points in cluster i and total number of points is n. Then* **cluster entropy** *E is equal to*

$$E = -\sum_{j=1}^{m} \frac{n_j}{n} \cdot \ln\left(\frac{n_j}{n}\right), \tag{2}$$

where $\frac{n_j}{n}$ is the proportion to the number of points in the cluster j.

Cluster entropy measures the level of concentration of points (independently of their importance) across clusters. High concentration leads to lower cluster entropy, whereas low concentration (i.e. more equal distribution of points across clusters) leads to higher cluster entropy.

Lemma: *If we have two clusters A and B with n_a and n_b number of points ($n_a < n_b$), and we are adding the point p then the maximal increase of cluster entropy happens if we add the point p to the smallest cluster.*

PROOF: Let $n_a + n_b = n$ and $\frac{n_a}{n+1} = t$ where $t < \frac{1}{2}$. Then $\frac{n_b}{n+1} = 1 - t$. When we add the point p to a cluster the proportion of points in this cluster is equal to $t + x$. To find where the increase of cluster entropy will be maximal, we have to take derivative of cluster entropy and set it equal to 0.

$$\frac{d(-((t+x)\cdot\ln(t+x) + (1-t-x)\cdot\ln(1-t-x)))}{dx} = \ln\frac{1-t-x}{t+x}$$

By definition $t + x \leqslant 1 - t - x$ and, therefore, $\frac{1-t-x}{t+x} \leqslant 1$ and $\ln\frac{1-t-x}{t+x} \leqslant 0$. Thus, the derivative is non-positive and it attains the maximum value 0 at the point, where $\ln\frac{1-t-x}{t+x} = 0$, therefore $\frac{1-t-x}{t+x} = 1$, $t + x = 1 - (t + x)$ and $t + x = \frac{1}{2}$. The latter equation shows that these clusters should have equal number of points, i.e. there should

be equal distribution. To get this result we have to assign any new points to the smallest cluster.
□

For example, let's consider space with two clusters: the first one with five points and the second one with seven points. The cluster entropy is equal to 0.6792. If we add a new point to the first cluster the cluster entropy will increase to 0.6902. But if we add a new point to the second cluster the cluster entropy will *decrease* to 0.6663.

Thus, the level of cluster entropy should serve as a constraint during granulation process. It should never exceed certain minimal level. The latter depends on total number of points, number of subtasks, etc. Thus the task of granulation in situations, where we need to maximize granular entropy, can be rewritten as optimization problem:

$$\begin{array}{ll} \text{Maximize} & GE \\ \text{Subject to constrain} & E > E_{min}. \end{array} \tag{3}$$

As we proceed with the granulation process, we face two main problems. The first problem is to get exact measure of change of granular entropy and cluster entropy if we add a new point to a particular granule (or change relevance of some point); the second problem is to leverage the process, i.e. to find the tool to fine-tune the distribution of points across granules. The first problem is addressed in the next section; the second problem in sections four and five.

3. Leverage levels of granular and cluster entropy

Theorem 1: *Let granular entropy in system, with k granules (with at least one point in each of them) and total granular relevance G, is equal to GE. Let a point be added to the granule j, that has granular relevance G_j, or a point in this granule j increases its relevance. Then the rate of change of granular entropy is equal to*

$$-\frac{1}{G} \cdot \left(GE + \ln\left(\frac{G_j}{G}\right)\right) \tag{4}$$

PROOF: If we add one point to the granule j then the differential will be equal to

$$\sum_{i/j} \frac{G_i}{G+\Delta} \ln\left(\frac{G_i}{G+\Delta}\right) + \frac{G_i+\Delta}{G+\Delta} \ln\left(\frac{G_i+\Delta}{G+\Delta}\right) - \left[\sum_{i/j} \frac{G_i}{G} \ln\left(\frac{G_i}{G}\right) + \frac{G_j}{G} \ln\left(\frac{G_j}{G}\right)\right]$$

and the derivative will be equal to

$$\lim_{\Delta\to 0}\left\{-\frac{1}{\Delta}\left[\sum_{i/j} \frac{G_i}{G+\Delta} \ln\left(\frac{G_i}{G+\Delta}\right) + \frac{G_j+\Delta}{G+\Delta} \ln\left(\frac{G_j+\Delta}{G+\Delta}\right) - \left(\sum_{i/j} \frac{G_i}{G} \ln\left(\frac{G_i}{G}\right) + \frac{G_j}{G} \ln\left(\frac{G_j}{G}\right)\right)\right]\right\} =$$

$$\lim_{\Delta\to 0}\left\{-\frac{1}{\Delta}\left[\sum_{i/j}\frac{G_i}{G+\Delta}\ln\left(\frac{G_i}{G+\Delta}\right)-\sum_{i/j}\frac{G_i}{G}\ln\left(\frac{G_i}{G}\right)+\frac{G_j+\Delta}{G+\Delta}\ln\left(\frac{G_j+\Delta}{G+\Delta}\right)-\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)\right]\right\}=$$

$$\lim_{\Delta\to 0}\left\{-\frac{1}{\Delta}\left[\sum_{i/j}\frac{G_i}{G+\Delta}\ln\left(\frac{G_i}{G+\Delta}\right)-\sum_{i/j}\frac{G_i}{G}\ln\left(\frac{G_i}{G}\right)\right]+\left(-\frac{1}{\Delta}\right)\left[\frac{G_j+\Delta}{G+\Delta}\ln\left(\frac{G_j+\Delta}{G+\Delta}\right)-\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)\right]\right\}=$$

$$\lim_{\Delta\to 0}\left\{-\frac{1}{\Delta}\left[\sum_{i/j}\frac{G_i}{G+\Delta}\ln\left(\frac{G_i}{G+\Delta}\right)-\sum_{i/j}\frac{G_i}{G}\ln\left(\frac{G_i}{G+\Delta}\right)+\sum_{i/j}\frac{G_i}{G}\ln\left(\frac{G_i}{G+\Delta}\right)-\sum_{i/j}\frac{G_i}{G}\ln\left(\frac{G_i}{G}\right)\right]+\right.$$

$$\left.+\left(-\frac{1}{\Delta}\right)\left[\frac{G_j+\Delta}{G+\Delta}\ln\left(\frac{G_j+\Delta}{G+\Delta}\right)-\frac{G_j+\Delta}{G+\Delta}\ln\left(\frac{G_j}{G}\right)+\frac{G_j+\Delta}{G+\Delta}\ln\left(\frac{G_j}{G}\right)-\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)\right]\right\}=$$

$$\lim_{\Delta\to 0}\left\{-\frac{1}{\Delta}\left[\sum_{i/j}\left(\frac{G_i}{G+\Delta}-\frac{G_i}{G}\right)\ln\left(\frac{G_i}{G+\Delta}\right)+\sum_{i/j}\frac{G_i}{G}\left[\ln\left(\frac{G_i}{G+\Delta}\right)-\ln\left(\frac{G_i}{G}\right)\right]\right]+\right.$$

$$\left.+\left(-\frac{1}{\Delta}\right)\left[\frac{G_j+\Delta}{G+\Delta}\left[\ln\left(\frac{G_j+\Delta}{G+\Delta}\right)-\ln\left(\frac{G_j}{G}\right)\right]+\left(\frac{G_j+\Delta}{G+\Delta}-\frac{G_j}{G}\right)\ln\left(\frac{G_j}{G}\right)\right]\right\}=$$

$$=\sum_{i/j}\frac{G_i}{G^2}\ln\left(\frac{G_i}{G}\right)-\lim_{\Delta\to 0}\left(-\frac{1}{\Delta}\right)\sum_{i/j}\frac{G_i}{G}\ln\left(\frac{G+\Delta}{G}\right)+\lim_{\Delta\to 0}\left\{\left(-\frac{1}{\Delta}\right)\left[\frac{G_j+\Delta}{G+\Delta}\ln\left(\frac{G_i+\Delta}{G+\Delta}\right)-\right.\right.$$

$$\left.\left.-\ln\left(\frac{G_j+\Delta}{G}\right)+\ln\left(\frac{G_j+\Delta}{G}\right)-\ln\left(\frac{G_j}{G}\right)\right]+\left(\frac{G_j+\Delta}{G+\Delta}-\frac{G_j+\Delta}{G}+\frac{G_j+\Delta}{G}-\frac{G_j}{G}\right)\ln\left(\frac{G_j+\Delta}{G+\Delta}\right)\right\}=$$

$$=-\frac{1}{G}\left[\sum_{i/j}\left(-\frac{G_i}{G}\ln\left(\frac{G_i}{G}\right)\right)-\sum_{i/j}\frac{G_i}{G}\right]-\frac{1}{G}\left[\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)+\frac{G_j}{G}-1+\ln\left(\frac{G_j}{G}\right)\right]=$$

$$=-\frac{1}{G}\left[-GE+1-1+\ln\left(\frac{G_j}{G}\right)\right]=-\frac{1}{n}\left[GE+\ln\left(\frac{G_j}{G}\right)\right]\qquad\square$$

We can view granular entropy of the whole system $GE=-\sum_{j=1}^{k}\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)$ as consisting of the sum of entropies $GE_j=-\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)$ of granules *within* the given system. At this point we need to find out the how we can achieve the largest increase and largest decrease of granular entropy of the system.

Let the number of granules be k. We can rewrite the above formula as

$$-\frac{1}{G}\left[GE+\ln\left(\frac{G_j}{G}\right)\right]=-\frac{k}{G}\left[\frac{GE}{G}+\frac{1}{k}\ln\left(\frac{G_j}{G}\right)\right]=$$

$$=-\frac{k}{G}\left[\overline{GE}+\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)-\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)+\frac{1}{k}\ln\left(\frac{G_j}{G}\right)\right]=$$

$$= -\frac{k}{n}\left[\overline{GE} - \left(-\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)\right) + \left(\frac{1}{k} - \frac{G_j}{G}\right)\ln\left(\frac{G_j}{G}\right)\right] = \tag{5}$$
$$= \frac{k}{n}\left[GE_j - \overline{GE} + \left(-\ln\left(\frac{G_j}{G}\right)\right)\left(\frac{1}{k} - \frac{G_j}{G}\right)\right]$$

where $\overline{GE}$ is average granular entropy, $GE_j = -\frac{G_j}{G}\ln\left(\frac{G_j}{G}\right)$ is granular entropy of the chosen granule j.

The latter equation points out that the change in granular entropy will be greatest, when both the difference between entropy of the chosen granule and average granular entropy and the difference between ratio of granular relevance in the chosen granule to total granular relevance and mean proportion, that is equal to $-\frac{1}{k}$, is greatest. The change in granular entropy will be smallest if these differences are close to zero, i.e. the chosen granule has entropy that is closest to an average.

To achieve more flexible control over the process of granulation we can order granules by values $-\frac{G_j}{G}\left(GE + \ln\left(\frac{G_j}{G}\right)\right)$, where GE is granular entropy and G_j is granular relevancy of the granule j. Then we control the direction of change of granular entropy by choosing corresponding granule to add new point. If we want that the change from the current level of granular entropy be minimal then we choose the granule j whose value $-\frac{G_j}{G}\left(GE + \ln\left(\frac{G_j}{G}\right)\right)$ is closest to zero. If we want the maximal increase in concentration (and, therefore, the maximal decrease of granule entropy), then we assign a new point p to the granule j with the largest negative value of $-\frac{G_j}{G}\left(GE + \ln\left(\frac{G_j}{G}\right)\right)$. If we want the maximal decrease in concentration (and, therefore, the maximal increase of granule entropy) then we assign new point p to granule j with largest positive value of $-\frac{G_j}{G}\left(GE + \ln\left(\frac{G_j}{G}\right)\right)$.

While solving the optimization problem, we can simultaneously leverage cluster entropy.

Theorem 2: *Let total cluster entropy in system with k clusters (with at least one point in each of them) and n points (with n sufficiently large) is equal to E. Let a point be added to the cluster j that has m points. Then the rate of change of cluster entropy is equal to*

$$-\frac{1}{n}\left(E + \ln\left(\frac{m}{n}\right)\right) \tag{6}$$

4. 3. T-norm and T-conorm operators and granular entropy

Suppose that there are m granules and each of them has at least one point. At this stage we need an algorithm according to which every point will be assigned to a particular granule. During this process we should take care that all granules form compact subspaces; the assignment of points to granules is made only in cases where compactness of granules is preserved.

As we define the distances between point p and granules, we get the leverage over the granulation process. The new point is added to the granule if the distance from this point to the granule is the smallest. We leverage the process by defining the way the distance between a point and granules is determined, and, therefore to what granule this point will be added.

Definition: *: Suppose we have a granule A with n points whose relevancies are equal to $d_1, d_2, \ldots, d_n$. Then* **granular distance** *D between point p that has relevancy d_{n+1} and granule A is computed as*

$$D = F(d_1, d_2, \ldots, d_n, d_{n+1}) - F(d_1, d_2, \ldots, d_n), \tag{7}$$

where F is monotonic, commutative, associative operator $[0, 1]^n \to [0, 1]$.

Initially we have m granules to which point p can be assigned, with granular distances at between point p and these granules equal to $D_1, D_2, \ldots, D_m$; we assign the point p to such granule i, that $D_i = \min_j (D_j)$.

For example, if the operator F is maximum operator and the relevancies in first and second granule are equal to $(0.2, 0.4, 0.7)$ and $(0.5, 0.6, 0.3)$ correspondingly, then the granular distances from the point with relevancy 0.45 are equal to 0.7 and 0.6 and the point is assigned to the second granule. The resulting operator, that determines to which granule the new point will be assigned, is minimax operator.

This operator F, whose arguments are points' relevancies, can serve as leveraging mechanism in the process of granulation.

What properties can be associated with this operator? First of all, the result of the computation should not depend on ordering of points in the granule, i.e. this operator should be commutative.

$$F(d_1, d_2) = F(d_2, d_1).$$

Second, we should be able to extend an operator from n to $n + 1$ arguments. Thus, this operator should be associative.

$$F(d_1, F(d_2, d_3) = F(F(d_1, d_2), d_3).$$

Third, it should be monotonic: an increase of relevance of some point in the granule from d_3 to d_2 should not lead to a decrease of granular distance between point p with relevance d_1 and this granule.

$$F(d_1, d_2) \geqslant F(d_1, d_3) \text{ if } d_2 \geqslant d_3.$$

Fourth, the more points in the granule the larger corresponding granular distance should be. Thus,

$$F(d_1, d_2, \ldots, d_k, d_{k+1}) \geqslant F(d_1, d_2, \ldots, d_k).$$

The inequality should hold also if the relevancy d_{k+1} is equal to 0. The latter case can happen if we add to the granule an additional point that happens to coincide with the point p. But

$$F(0,0) = F(0) = 0.$$

And therefore

$$F(d,0) = F(d, F(0,0)) = F(F(d,0),0).$$

From the latter equation follows that $F(d,0) = d$. Thus, we have additional requirement: an operator F has identity 0, i.e. $F(d,0) = d$.

Operators, that satisfy all four properties - commutativity, associativity, monotonicity, and zero identity - are called t-conorm operators. Foremost among them is the maximum operator. If we apply the maximum operator for evaluation of granular distances, the resulting granulation tends to be more equal. But it is not necessary. If we have two granules with relevancies equal to $(0.25, 0.3, 0.35, 0.45)$ and $(0.5, 0.6)$ the granular distances and the point with relevancy 0.4 are equal to 0.45 and 0.6 correspondingly, and the point p will be assigned to the first granule even though it already has the largest number of points.

What we need is operator that takes into account not just the granular distances from point p to points in a granule but the number of points in a granule as well. For this purpose we can use other t-conorm operators. For example, if we apply Lucaŝevich t-conorm $S(a,b) = a + b - ab$ in the above example then the granular distances will be 0.8874 and 0.88 and point p will be assigned to the second granule, even though all relevancies in the first granule are less then relevancies to the points in the second granule. On other side, if the first granule consisted of just three points with relevancies equal to $(0.25, 0.3, 0.35)$, then the granular distance to point p (calculated as Lucaŝevich t-conorm) is equal to 0.6587. The latter is less then granular distance from the point p to second granule (0.88) and, therefore, point p will be assigned to the first granule, even though it has more points.

Thus, choosing the appropriate t-conorm operator, we can approximate the desired granular entropy (because values of granule entropy are discrete, we can not organize granulation process to get exact value). And, as the process of adding new points is iterative, we can make sure that difference in size of granules will be within desired bounds.

If, on other hand, we want to get minimal granular entropy, we have to assign new points to existing granules with the largest number of points. As a result most of points will be concentrated in one granule (because we have fixed the number of granules with at least one point in each of them, granular entropy will never reach 0). In such a case the larger the number of points in the granule the smaller granular distance between the point p and this granule is. Therefore, the granular distance between the point p and some granule should not increase if we add another point to the granule. Thus,

$$F(d_1, d_2, \ldots, d_k, d_{k+1}) \leqslant F(d_1, d_2, \ldots, d_k).$$

This inequality should also hold if we add point with the maximal relevancy (which is equal to 1). But $F(1,1) = F(1) = 1$. Therefore, we have

$$F(d,1) = F(d, F(1,1)) = F(F(d,1),1).$$

From the latter equation follows that identity equal to 1. Operators, that satisfy four properties - commutativity, associativity, monotonicity, and one identity - are called t-norm operators. The most common t-norm operator is the minimum operator.

We can get control of granulation by using the t-conorm operator

$$T^{\star}(x,y) = \log_s \frac{(s^x - 1)(s^y - 1)}{s-1} \tag{8}$$

and its dual t-conorm operator

$$S^{\star}(x,y) = 1 - T^{\star}(1-x, 1-y) \tag{9}$$

by varying the parameter s: if we want to add a new point to the granule with largest number of points then we use t-norm $T^{\star}$ with large s. On the other side if we want to add the point p to the granule with the least number of points we use t-conorm $S^{\star}$.

5. 4. Uni-norm operators and leveraging of granulation process

Operators that determine granular distance should satisfy at least three properties: commutativity, associativity, and monotonicity. If we need as a result of granulation process to get as large granular distance as possible we use zero identity; if we want as small granular distance as possible we use one identity. If we want to leverage granular entropy toward some particular value we should use identity in between zero and one.

Definition: *Operators* $U : [0,1] \times [0,1] \to [0,1]$ *that satisfy four conditions:*

a) *Monotonicity*
b) *Commutativity* $U(d_1, d_2) = U(d_2, d_1)$
c) *Associativity* $U(d_1, U(d_2, d_3)) = U(U(d_1, d_2), d_3)$
d) *Having fixed identity* e *i.e.* $U(d,e) = d$ *for any* d *are called* **uni-norm** *operators.*

We can construct uni-norm from any t-norm operator T and t-conorm operator S:

$$\begin{aligned} &\text{If } x,y \in [0,e] \text{ then } U(x,y) = e \cdot T\left(\frac{x}{e}, \frac{y}{e}\right) \\ &\text{If } x,y \in [e,1] \text{ then } U(x,y) = e + (1-e)S\left(\frac{1-x}{1-e}, \frac{1-y}{1-e}\right) \\ &\text{If } x \in [0,e], y \in [e,1] \text{ then } U(x,y) = \min(x,y) \text{ or } U(x,y) = \max(x,y) \end{aligned} \tag{10}$$

For example, if $T(x,y) = \min(x,y)$ and $S(x,y) = \max(x,y)$ then

$$U_1(x,y) = \max(x,y), \text{ if } x,y \in [e,1] \tag{11}$$

$$U_1(x, y) = \min(x, y), \text{ otherwise}.$$

Theorem 3: *If granular distance is defined by an operator U_1, then granular entropy is monotonically decreasing on e.*

PROOF: Let identity in the operator U_1 be equal to e_1, and identity of operator U_2 be equal to $e_2 > e_1$.

If $x, y \in [e_2, 1]$
$$U_1(x, y) = U_2(x, y) = \max(x, y).$$
If $x \in [0, e_1]$
$$U_1(x, y) = U_2(x, y) = \min(x, y).$$
If $x \in [e_1, e2], \ \ y \in [0, e_1]$
$$U_1(x, y) = U_2(x, y) = \min(x, y).$$
If $x \in [e_1, e_2], \ \ y \in [e_1, 1]$
$$U_1(x, y) = \max(x, y), \ \ U_2(x, y) = \min(x, y)$$
and $U_1(x, y) \geqslant U_2(x, y)$.

Therefore, $U_1(x, y) \geqslant U_2(x, y)$ for all x, y. □

As we increase the value of identity e the granular distance D is decreasing (non-increasing). Therefore, granular entropy is decreasing (non-increasing). For example, let the relevancies of points in granules A, B and C be equal to $(0.2, 0.3, 0.9)$, $(0.4, 0.6, 0.7)$ and $(0.1, 0.5, 0.8)$ correspondingly. If e is equal to 0.5, then the granular distances from point p with relevancy 0.6 to the granules A, B and C are equal to 0.2, 0.4, 0.1 and the point p will be assigned to the third granule. If identity e is equal to 0.1 then the granular distances are 0.9, 0.7, 0.8 and the point p will be assigned to the second granule. Thus, to achieve specific results we can fine-tune the identity level of uni-norm operator (not necessarily U_1) and get the desired level of granular entropy.

In the same example, we can look at the level of change of granular entropy (Theorem 2). Their values of $-\frac{1}{G}\left(GE + \ln\left(\frac{G_j}{D}\right)\right)$, where GE is granular entropy, G_j is granular relevancy of the granule j and G is total granular relevancy) are equal to 0.0163, -0.0268 and 0.0163, while their granular relevancy is equal 1.4, 1.7 and 1.4 correspondingly. Granular entropy of the whole system is equal to 1.0041. To get maximum increase of cluster entropy we have to add a new point to the fist cluster; to get maximum decrease of cluster entropy we have to add a new point to the third cluster; and make the minimal change in cluster entropy we have to add a new point to the second cluster.

6. Conclusion

In this paper we introduce the concepts of granular entropy and cluster entropy and describe how we can use them to represent granulation process as solution of optimization problem. To this end the theorem, that shows how the change in relevance of points in a granule affects granular entropy, was proved. To get additional leverage over granulation process we introduce the concept of granular distance and show how this leverage can be achieved by applying t-conorm and uni-norm operators to control granular distances and granular entropy.

References

[1] J.C. FODOR, R.R. YAGER and A. RYBALOV: Structure of uni-norms. *Int. J. of Uncertainty, Fuzziness and Knowledge-Based Systems*, **5** (1997), 411-427.

[2] C.J. HARRIS C.G. MOORE and M. BROWN: Intelligent Control - Aspects of Fuzzy Logic and Neural Nets. World Scientific, Singapore, 1993.

[3] W. PEDRYCZ: Computational Intelligence: An Introduction. CRC Press, Boca Baton, 1997.

[4] A. RYBALOV and R.R. YAGER: Control Clustering, Uni-Norm Operators and OWA operators. *IFSA/NAFIPS Conf.*, Vancouver, (2001).

[5] R.R. YAGER: On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. on Systems, Man and Cybernetics,* **18** (1998), 183-190.

[6] R.R. YAGER: Intelligent control of the hierarchical agglomerative granulation process. *IEEE Trans. on Systems, Man and Cybernetics*: Part B, **30** (2000), 835-845.

[7] R.R. YAGER and J. KACPRZYK: The Ordered Weighted Averaging Operators: Theory and Applications. Kluwer: Norwell, 1997.

[8] R.R. YAGER and A. RYBALOV: Full reinforcement operators in aggregation techniques. *IEEE Trans. on Systems, Man and Cybernetics*, **28** (1998), 757-769.

[9] R.R. YAGER and A. RYBALOV: Uninorm aggregation operators. *Fuzzy Sets and Systems*, **80** (1996), 111-120.

[10] L.A. ZADEH: Fuzzy sets and information granularity. In M.M. Gupta, R.K. Ragade, R.R. Yager (Eds), *Advances in Fuzzy Set Theory and Applications,* North Holland, Amsterdam, 1979.

[11] L.A. ZADEH: Fuzzy logic = Computing with words. *IEEE Trans. on Fuzzy Systems*, **4**(2) (1996), 103-111.

[12] L.A. ZADEH: Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy Sets and Systems*, **90** (1997), 111-117.