

Impact of familiarity on information complexity in human-computer interfaces

Maxim Bakaev

Economic Informatics dept., Novosibirsk State Technical University, Novosibirsk, Russia

Abstract. A quantitative measure of information complexity remains very much desirable in HCI field, since it may aid in optimization of user interfaces, especially in human-computer systems for controlling complex objects. Our paper is dedicated to exploration of subjective (subject-depended) aspect of the complexity, conceptualized as information familiarity. Although research of familiarity in human cognition and behaviour is done in several fields, the accepted models in HCI, such as Human Processor or Hick-Hyman's law do not generally consider this issue. In our experimental study the subjects performed search and selection of digits and letters, whose familiarity was conceptualized as frequency of occurrence in numbers and texts. The analysis showed significant effect of information familiarity on selection time and throughput in regression models, although the R^2 values were somehow low. Still, we hope that our results might aid in quantification of information complexity and its further application for optimizing interaction in human-machine systems.

1 Introduction

Human operators so far remain an indispensable part of most control systems, receiving, processing, and outputting information just like any other its component. Although the recognized advantages of humans' in comparison to machines are their integrity, adaptability, and self-development capability, they also remain a somehow limiting factor in semi-automated control progress [1]. Probably the main reason for this is the unresolved problem of organizing optimal interaction between humans and computers (HCI). In particular, the cognitive aspect of HCI remains quite less explored compared to physical (ergonomic) one, as the work of human brain and memory remain somehow of a mystery to the present day.

Naturally, it would be desirable to develop theories describing how humans perform information processing, what are their productivity and constraints in that, so that complexity of user interfaces could be reasonably minimized, which is generally considered to be desirable [2]. In theoretic information science there exist such fields as Computational complexity theory and Analysis of algorithms, but their application in HCI at the moment seems limited. Indeed, complexity in interfaces is more commonly explored on a qualitative level, at best measured as *chunks of information*, per the infamous Miller 7 ± 2 law for working memory capacity. It should be said that the law indeed is an interesting fundamental regularity, which is seemingly applicable in software interfaces as well – e.g. in [3] the recommended number

of interactions per component in a component-based system is 5.

As for more human-oriented research, [4] compares effect of complexity (enhancements in software functionality) for elder and regular computer users. An extensive review on complexity is provided in [5], with the emphasis on organization of nuclear plants control rooms. A set of measures for information complexity that can be used to evaluate automation systems is presented in [1], where the authors view it as the combination of three basic factors: numeric size, variety, and relation. The stages of brain information processing are identified as perception, cognition, and action. In [6] they propose operation, transition, and screen complexity, and the quantitative interface complexity measures are based on informational entropy (H) concept, or the information quantity. Quite a straightforward formula:

$$T = 0.03 * H, \quad (1)$$

where T is time required for receiving and processing information, was mentioned in [7, ref 84], but we feel the constant it employs it is at best averaged value, while each user, task, and concept should be special in HCI.

We also believe that currently limited applicability of Computational complexity theory in HCI is due to significant difference of users, even within a user group. In our current paper we seek to explore the subjective side of the information complexity, by analyzing the impact of *familiarity* (well-known versus unfamiliar information) in human information processing. It should be mentioned that there are clearly different levels for complexity and familiarity in user interfaces, the structure of which so far remains ambiguous. For example, we

would identify them as *Task (Functionality) – Thesaurus – Layout – Material – Behaviour*, and so far we focus on the Material level, exploring the impact of familiarity in interface elements (the level corresponding to icons, buttons and other controls, etc.). So, in Section 2 we outline some related methods, while in Section 3 we describe the experiment conducted with real subjects and analyze its implications for the research problem.

2 Methods

Familiarity is quite extensively studied in such domains as marketing (product and brand familiarity), sociology, psychology, medicine [8], etc. Although in HCI the studies of familiarity started long before the era of WWW (e.g. [9]), currently the number of corresponding works is rather limited. In a study of effect of user familiarity with a website on its perceived privacy [10], the authors described cognitive and affective functions of familiarity. Among the more interface-related research we'd like to note [11], where the authors employ the concept of cognitive distance, which is important since it expresses the familiarity in a quantitative way. So, next we consider some other approaches and findings that use operable quantitative measures.

2.1 Human information input and output

Human, as information processor, can input and output information, and for both processes measures of speed can be introduced (R_{in} and R_{out}), as information volume time derivative (or, we may also say, second derivative of knowledge). Naturally, absolute values of the Rs depend upon measurement units, and this adds certain confusion when speed estimations from different research works are measured. All in all, sources seem to agree that the upper limit of human processor to consciously perceive information is $R_{in} < 10^2$ bit/s, although human nervous system may process as much as $R_{in} = 10^9$ bit/s subconsciously; this fact probably being one of the grounds for popular talks about hidden reserves of human mind.

As for information output, it is said that maximum speed of human's vocal cords operation matches up to $R_{out} = 5$ syllables/s in medium-term, which corresponds to $R_{out} = 12.5$ sounds/s for English, while for other languages this value may differ. For humans equipped with instruments or tools, the values are generally presented in bits/s or chars/s. Hand-writing would be limited at 3.5 char/s, while average typing speed is about 6.7 char/s, with 10 char/s being the top record, which still seems to be inferior to oral speaking. In any case, maximum speed of human processor as information transmitter is estimated at 40 bit/s [7, pp. 21-22].

The values for conscious information reception seem to be only marginally higher. R_{in} for speaking comprehension is estimated at 3-4 words/s, for reading comprehension – at 2.5 words/s or 18 bit/s, for reading aloud – 30 bits/s, for speedy reading (Russian language) – at up to 11.7 words/s or 80 bit/s. Most sources agree that although general R_{in} for reading is at about 45 bit/s, the

capability of human visual perception for long-term work is at about 8 bit/s only [7, p. 23].

The speed can be increased to certain extent by using multi-dimensional signals instead of simple uniform ones – e.g. tone and volume instead of just volume; or size, brightness, and colour, instead of just one varying visual stimulus – but the effect is not additive per each introduced dimension. The ability to excel with signals of increased complexity somehow explains why humans are notably good, in comparison to machines, at image recognition. In one of our previous experiments we found that the higher was the diversity of objects in visual search/selection tasks, the higher throughput (TPS) was demonstrated – from 42 bit/s for objects' vocabulary size = 2 (simplest tasks) to 78 bit/s for vocabulary size = 29 (Russian letters) [12]. It should be noted that recognition speed for familiar letters and digits was previously estimated as 55 bit/s [7, ref 75].

Another factor that is known to increase the speed is training, which in particular allows operators to perceive information on a higher level, processing fewer numbers of chunks/elements. For example, while novice pilots need to individually check every display, expert ones require only a glance on the familiar console to decide if the controlled object, i.e. airplane, is in order. For relatively simple case of incoming textual information flow, it may be represented by introducing the *coefficient of new information density* [7, p. 144-145].

This calls for accounting the degree of information familiarity for a recipient – a qualitative and subjective aspect, and they are generally not convincingly considered in Information Theory-related models.

2.2 Human Processor Model

The Human Processor Model (also MHP) that was initially developed in the 80s is essentially a cognitive modeling method predicting cognitive and motor processing times. In this it is based on pre-defined experimental values, and its general structure, which draws similarities between humans and modern computers' architecture, can be outlined as shown on Fig. 1.

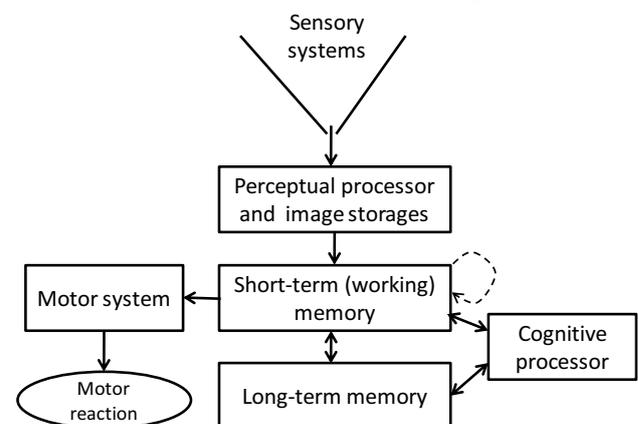


Figure 1. Principal components of the Human Processor Model.

It can be seen that the cognitive sub-system (incorporating working memory, long-term memory, and cognitive processor) is the most interlinked and indeed

the most complex one. So, although the experimental statistics are quite reliable for senses (e.g. storage time for vision is 0.3-1 s, for hearing is up to 3.5 s), and very sound model exists for the motor system (Fitts' law), the cognitive aspect is quite less developed. The alleged *cognitive processor cycle time* varies in the range of 25-170 ms, but it remains rather unclear how to determine the number of cycles, in which cases information exchange with long-term memory starts and how exactly, what are the particulars for various user groups (e.g. senior people) and contexts of use (e.g. fatigue), etc. Thus we believe the cognitive aspect requires further exploration.

2.3 The Hick-Hyman law

Once, W.E. Hick noted that reaction time (RT) when choosing from N equally probable alternatives is proportional to the logarithm of their number:

$$RT \sim k * \log_2(N + 1), \quad (2)$$

where k is the rate of gain of information. Afterwards, R. Hyman showed that RT in fact is linearly related to information quantity, i.e. the entropy of the set of stimulus (H_T):

$$RT = a_H + b_H * H_T, \quad (3)$$

where a_H and b_H are empirically defined constants. The slope in thus formulated Hick-Hyman law (1), b_H , in simplest cases is believed to be equal to 150 ms [13], then the corresponding Hick's rate of gain of information (b_H^{-1}) is equal to 6.7 bits/sec. This rate varies significantly due to subjects' age, gender, arousal, fatigue, etc., and also due to training a familiarity with the stimuli. So, it has been shown e.g. by [13] that for unknown alternatives RT increases even linearly with N, but has almost no increase for well-known options – which also calls for introduction of the familiarity factor in the corresponding model. Thus, in the next section we explore the impact of familiarity in the specially designed experiment and outline some notable results.

3 The experiment description

3.1 Subjects

Our experimental investigation was performed in several sessions during two years, and the subjects were recruited among Master students of Web Design major of a university located in Russia. In total, 33 subjects took part in the experiment, 14 of them being female and 19 being male. Their age ranged from 20 to 42 ($M=24.5$, $SD=5.21$). All subjects had normal or corrected to normal vision and high level of experience in using computers and mouse. Before the experiment, data regarding the participants' age and gender were gathered. All subjects participated in the experiment voluntarily, and prior to the experimentation informed consents were obtained. Each subject performed non-recordable test run of trials with random combinations of factors until fully mastering the tasks, to negate the effect of practice. The sessions with three different groups of participants took place with

1-year intervals on same or similar computer equipment (mouse, etc.) and screen sizes and resolutions.

3.2 The experiment design and procedure

In the related part of the experiment (more detailed description of the whole experiment can be found in [12]), the subjects were asked to undertake selection tasks. They were presented with a starting square positioned randomly (to negate the effect of movement direction) next to the work area of varying size S_0 (256x256, 512x320, 512x512, or 1024x640 pixels), in which N (8, 32, or 128) objects of varying size W (16, 32, or 64 pixels) would appear.

The participants were tasked to click the starting position with a mouse pointer and then, "as fast and as accurately as possible", find the target, move the mouse pointer and click the target – the coordinates of both clicks were recorded. If the second click was outside the target, error was recorded: the participant was taken to a next trial, but if overall error rate exceeded 10%, the message asking to increase accuracy would be displayed. The dependent variables were selection time (ST, between the two clicks) and error rate (E). Errors emerged when clicks were made outside the sought target (i.e. a false alternative was selected or the target just missed), and in case of error rate exceeding rate 10%, the message asking to increase accuracy was displayed between trials.

To measure and record the values of independent and dependent variables, an online application was developed with PHP and MySQL. The subjects used two web browsers, Chrome and Firefox, to negate the possible representation effect, and the performance time was measured with JavaScript to eliminate any server-side delay. Although this measurement method is known to have certain disadvantages, we still considered it appropriate for our rather exploratory study.

3.3 The target objects

Since our goals included exploring of the effect of vocabulary and familiarity, we employed different types of objects as the targets: geometric shapes, numbers or Cyrillic letters. The sizes of vocabulary K, corresponding to the number of various possible objects in the type, were 2, 10, and 29 respectively (we excluded several Cyrillic letters that looked too similarly to each other).

Since virtually all our subjects had Russian as their native language, for the Cyrillic letters we used frequencies based on the Russian National Corpus [14] (with the adjustment for the frequencies of the 4 letters excluded from the experiment), see Table 1.

For the digits, strictly speaking, the frequencies would correspond to uniform distribution (10% each), but only for long enough numbers. In our case, however, all the corresponding objects were represented as a single digit, so we chose to use the well-known Benford's law formula for the first digit [15]. We also had the consideration that the subjects likely to pay attention to initial (left-side) digits more often in their life, so their familiarity should follow the first-digit law more closely

than the uniform distribution. The resulting frequencies for digits (with zero excluded from the analysis) are shown Table 2; and we'll be using percentage values rather than decimals, for better readability.

Table 1. Observed frequencies for the 29 Russian letters.

Rank	Letter	Freq.	Rank	Letter	Freq.
1	О	11.2%	16	Я	2.0%
2	Е	8.6%	17	Ы	1.9%
3	А	8.1%	18	Ь	1.8%
4	И	7.5%	19	Г	1.7%
5	Н	6.8%	20	З	1.7%
6	Т	6.4%	21	Б	1.6%
7	С	5.6%	22	Ч	1.5%
8	Р	4.8%	23	Х	1.0%
9	В	4.6%	24	Ж	1.0%
10	Л	4.4%	25	Ш	0.7%
11	К	3.5%	26	Ю	0.6%
12	М	3.3%	27	Ц	0.5%
13	Д	3.0%	28	Э	0.3%
14	П	2.9%	29	Ф	0.3%
15	У	2.7%			

Table 2. Presumed frequencies for the digits.

Digit	Frequency
1	30.1%
2	17.6%
3	12.5%
4	9.7%
5	7.9%
6	6.7%
7	5.8%
8	5.1%
9	4.6%

4 The experiment results

4.1 Descriptive statistics

In the experiment, data for 6231 trials were collected, of which 5832 (93.6%) were considered valid (excluded were the ones when subject made erroneous double-click at start, or when the ST exceeded 15000 ms). Table 3 contains mean selection times (ST) and error levels (E) per number of targets (N) and vocabulary sizes (K).

Table 3. Mean ST (ms) and E (%) per N and K.

N	K	2	10	29	Mean ST, E
8	964	1477	1456	1298	6.1 5.4 3.9 5.1

32	1043 4.2	2485 3.0	2710 3.7	2080 3.6
128	1411 4.7	5003 7.3	5570 8.2	3853 6.7
Mean ST, E	1107 5.1	2688 5.0	2839 4.9	2196 5.0

4.2 Regression analysis

In our previous analysis of the experimental data [12], we first built regression for the classic Hick's law, which turned out to be significant ($F_{1,5830}=1494$; $p<0.001$), but had quite poor $R^2=0.204$ and perplexing negative intercept:

$$ST = -691 + 616 * \log(N) \quad (4)$$

Among several factors potentially reflecting the complexity of interface that we tried in the regression, the "effective interface capacity", defined as number of alternatives multiplied by logarithm of their vocabulary size, showed the best fit:

$$ST = 1123 + 7.92 * N * \log(K) \quad (5)$$

The model showed high significance ($F_{1,5830}=4160$; $p<0.001$) and had relatively high $R^2=0.416$, compared to the $R^2=0.666$ for Fitts' law model (movement tasks) that we obtained in the control part of the experiment.

In accordance with our current goal, we added frequency as factor to the model. Digits and letters were considered separately, since their frequencies (F_D and F_L respectively, expressed in percentages) are valid only relatively to the same characters:

$$ST_D = 1691 + 8.7 * N * \log(K) - 25.3 * F_D \quad (6)$$

$$ST_L = 1469 + 6.89 * N * \log(K) - 27.2 * F_L \quad (7)$$

The model (6) was highly significant ($F_{2,1734}=453$; $p<0.001$), with $R^2=0.343$, and the F_D factor was also significant ($p<0.001$). The model (7) was also highly significant ($F_{2,1900}=587$; $p<0.001$), with $R^2=0.382$, but the F_L factor was significant only at $\alpha=0.1$ ($p=0.094$).

4.3 Throughput

In our previous work [12] we introduced the concept of throughput for selection tasks (TPS), with Fitts' index of task difficulty substituted for the "effective interface capacity" factor:

$$TPS = \frac{1}{Z} \sum_{i=1}^Z \left(\frac{1}{M} \sum_{j=1}^M \frac{N * \log(K)_{ij}}{ST_{ij}} \right). \quad (8)$$

We then showed reasonable validity of the TPS concept and discovered that K had significant effect on TPS, which showed highly significant effect (in ANOVA, $F_{2,5829}=33.1$; $p<0.001$). Mean TPS values for K=2, 10, and 29 were 41.9, 58.6, and 77.6 bit/s respectively. We reasoned that higher values of K led to higher TPS since higher diversity of objects reflect their higher complexity, and humans are prominent at recognizing complex images, rather than quickly and accurately performing simple and typical tasks.

To explore the effect of familiarity on TPS, we built regression models for digits and letters with respective familiarities as the factor:

$$TPS_D = 50.1 + 0.725 * F_D \quad (9)$$

$$TPS_L = 72.5 + 1.487 * F_L \quad (10)$$

The model (9) was highly significant ($F_{1,1735}=16.1$; $p<0.001$), and the F_D factor was also significant ($p<0.001$), although the R^2 was very low, at 0.01. The model (10) was also highly significant ($F_{1,1901}=5.33$; $p<0.001$), the F_L factor was significant only at $\alpha=0.05$ ($p=0.021$), and the R^2 was even lower, at 0.003. As a comparison, in [12] the subjects' age factor had somehow comparable effect on their TPS ($F_{1,5830}=9.48$; $p=0.002$; $R^2=0.002$).

5 Conclusions

Controlling information complexity when designing interfaces in human-machine systems remains a crucial, but unresolved issue, due to lack of agreed quantitative metrics. Effective interface would be one that only presents relevant properties of the controlled object, and the final goal is to allow prompt and minimal-error interaction, and information load on human operators must be not maximized, but optimised according to their processing capabilities. However, there is still ambiguity regarding the ways to measure and optimize information complexity in HCI, in particularly because of its seemingly subject-specific nature.

In our paper we explored the impact of information familiarity on perceived complexity, having compared several existing approaches and finding none satisfactory in this regard. In the analysis of our experimental data, we based on our previous work [12], in which we formulated the concept of selection tasks throughput (8) that utilizes the "interface information complexity" (5) analogously to Fitts' index of difficulty. The results suggest that familiarity of digits and letters sought by the 33 subjects in the experiment did have significant effect on selection time (6), (7). The regression models for TPS (9), (10) also showed that familiarity was significant, although R^2 remained quite low.

The values obtained for TPS in the experiment, 42-78 bit/s, seem reasonable, as recognition speed for familiar letters and digits was previously estimated as 55 bit/s [7]. We see limitations of our study in its rather exploratory nature, which in particular lead to somehow low R^2 values. Also, in terms in conceptual validity we would agree that it's not so far satisfactory justified that frequency of digits' and letters' occurrence exactly correspond their familiarity for subjects. Still, we hope that our results might aid in defining the quantitative measure of information complexity and its further application for optimizing interaction in human-machine systems.

Acknowledgements

The reported study was funded by RFBR according to the research project No. 16-37-00184 mol_a.

References

1. J. Xing. Measures of information complexity and the implications for automation design. No. DOT/FAA/AM-04/17. Fed. Aviation Administration Oklahoma City Ok Civil Aeromedical Inst (2004).
2. A. Holzinger et al. On complexity reduction of user interfaces for safety-critical systems. Multidisciplinary Research and Practice for Information Systems, pp. 108-122 (2012).
3. U. Kumari and S. Upadhyaya. An interface complexity measure for component-based software systems. Int. J. of Com. Applications, 36.1 (2011).
4. Wu, Yu, and C. Slyke. Interface complexity and elderly users: revisited. Proc. SAIS, No. 51 (2005).
5. M.L. Cummings et al. Human-system interface complexity and opacity part i: literature review. Mass. Inst. of Tech., Cambridge, MA (2010).
6. H.G. Kang and P.H. Seong. An information theory-based approach for quantitative evaluation of user interface complexity. Nuclear Science, IEEE Transactions on 45.6, pp. 3165-3174 (1998).
7. V.F. Prisnyakov, L.M. Prisnyakova. Mathematic modelling of information processing by human-machine systems' operator (in Russian). Moscow: Mashinostroenie (1990).
8. L.A. Libby et al. Recollection and familiarity in schizophrenia: a quantitative review. Biological psychiatry 73.10, pp. 944-950 (2013).
9. B.L. Somberg and M.C. Picardi. Locus of the information familiarity effect in the search of computer menus. Proc. of the Human Factors and Ergonomics Society, 27 (9), SAGE (1983).
10. Y. Li. The impact of disposition to privacy, website reputation and website familiarity on information privacy concerns. Decision Support Systems 57, pp. 343-354 (2014).
11. C. Gough, R. Green, and M. Billingham. Accounting for user familiarity in user interfaces. Proc. of the 7th ACM SIGCHI New Zealand chapter's international conference on CHI (2006).
12. M. Bakaev, T. Avdeenko. A Quantitative Measure for Information Transfer in Human-Machine Control Systems. Proc. IEEE Int. Siberian Conf. on Control and Communications (SIBCON), IEEE (2015).
13. L.E. Longstreth et al. Exceptions to Hick's law: Explorations with a response duration measure. J. of Exp. Psychology: General, 114, pp. 417-434 (1985)
14. Wikipedia entry on Частотность (Frequency) <https://ru.wikipedia.org/wiki/Частотность>, 21.03.2016.
15. Wikipedia entry on Benford's law https://en.wikipedia.org/wiki/Benford's_law, 21.03.2016.