Counting Effective Number of Buttons: An Informational Analysis of Input Device Performance

Toru Nakata

Digital Human Research Center, National Institute of Advanced Industrial Science and Technology
Aomi 2-41-6, Koto, Tokyo, 135-0064, JAPAN
toru-nakata@aist.go.jp

Abstract

A quantitative method to evaluate theoretical efficiency of input devices is proposed. Using informatics, the method answers effective degrees of freedom in input devices; it evaluates how many buttons are efficiently used for operation. Comparing traditional methods represented by "magical number of seven," this method is versatile for general usability tests of any input devices, since the method requires only frequencies of stimuli and responses between machines and users.

1 Introduction: Proper number of buttons is always seven?

Degrees of freedom (DoF) of input devices cause inefficient operations, when it is excessive or insufficient. Excessive buttons and inadequately narrow menu trees make users tired and irritated and erroneous (Norman, 1988). Designers of input devices must estimate proper DoF.

Theoretical definition of DoF of input devices is the number of input devices' internal states except idling states. For instance, an input device consisting of three radio buttons (i.e. exclusive buttons) has three DoF. If an input device has an idling state and three buttons that are non-exclusive and selectable at the same time, the DoF is seven. In menu trees, DoF definition becomes more complicated. DoF of a decision tree equals to its information uncertainty (see Huffman, 1952).

In conventional studies, proper DoF values of input devices were estimated empirically. The most famous "proper" DoF is so-called "magical number of seven" (Miller, 1956). Seven (plus minus two) has been considered as the universally proper (or at least recommendable) DoF number for any user, any machine, and any condition. The value of seven was derived from Miller's experimental results on measuring human ability of short-term memory.

Obviously seven or other fixed numbers are not always good for any case. Cognitive abilities of users and difficulty of tasks vary largely. But still in 1997, ISO 9241-14 recommends seven as number of menu items.

Proper DoFs for various input devices are measurable by using informatics theory. This paper reports a method to evaluate DoF, referring an example experiment.

2 Method to estimate number of effective buttons

2.1 Target of evaluation: Information transfer

The method proposed in this paper considers desirable input devices as input device structures that realize maximum information transfer from the users to the machines.

Maximum information transfer is realized in case of minimum error occurrence, minimum key strokes, and minimum number of operation steps. Users can operate machines with the least cognitive and physical efforts.

Amount of information transfer depends on DoF of input devices, ability of users, and other conditions.

2.2 Evaluation framework: Three phases of human machine interaction

We must clarify flow of information during human machine interaction in order to evaluate pure information transfer without complicated factors.

Process of inputting a command to machines consists of the following three phases:
A) Cue. The machine or environment emit signals to the user.
B) Decision. The user decides what to do.
C) Operation. The user makes body action to operate the machine.

Process from A to B is interpretation by the user. Since information transfer in the A—B process depends mostly on goodness of cue representations or difficulty of decisions, we should treat it separated from usability of the input device.

Transition from B to C is process of handling of the input device. Usability of the input device appears in the B—C process, so we focus on this step.

2.3 Calculation of information transfer amount

Garner and Hake (1951) developed the method to calculate amount of information transfer in the B-C process. Fitts (1953) applied the method for usability testing; he evaluated information transfer from pilots to control stick of air-planes. Ironically, this method was born before magical number of seven. The reason that this method has been not famous may be that it requires a certain amount of calculation. Also Fitts stopped his analysis at calculation of information transfer amount, T, without interpreting practical meaning of T.

To review Fitts's method, we denote some symbols:

B_i : Operational action planned by the user at step B. (i = 1, …, m). (1)

C_j : Operational action carried out by the user at step C. (j = 1, …, n). (2)

f_{ij} : Occurrence number of C_j after the user selected B_i . (3)

$$N \equiv \sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}$$: Total number of interaction. (4)

$$p_{ij} \equiv \frac{f_{ij}}{N}$$: Probability that pair B_i and C_j takes place. (5)

$$p_{i\bullet} \equiv \sum_{i=1}^{m} p_{ij}$$: Probability that B_i is selected. (6)

$$p_{\bullet j} \equiv \sum_{j=1}^{n} p_{ij}$$: Probability that C_j is carried out. (7)

We got the following informational uncertainties or entropies, assuming $0 \log 0 = 0$ in the calculations.

$$H(B) \equiv -\sum_{i=1}^{m} p_{i\bullet} \log_2 p_{i\bullet}$$: Informational uncertainty of B_i selection. (8)

Direct meaning of H(B) is amount of information contained in user's single decision. H(B) may interpret as difficulty of the task increased by change of environments. Depending the experiment, other interpretation may be suitable; H(B) may represent characteristic of users.

$$H(C) \equiv -\sum_{j=1}^{n} p_{\bullet j} \log_2 p_{\bullet j}$$ Informational uncertainty of C_j selection. (9)

H(C) is variety of user's behaviour.

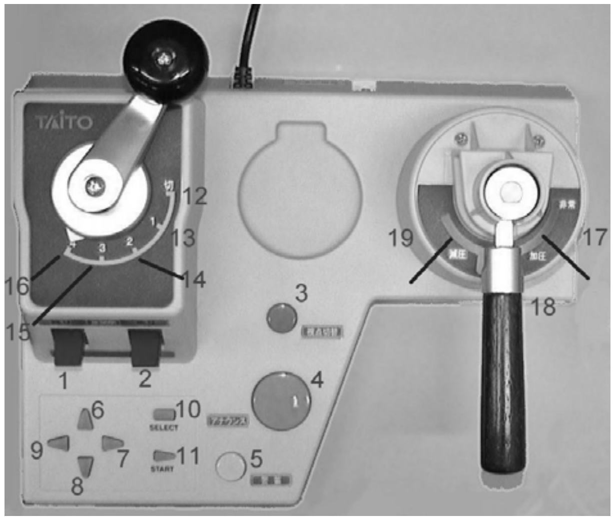

Fig. 1. Experimental input device.

$$H(B \to C) \equiv -\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij} \log_2 p_{ij} \quad \text{: Informational uncertainty of B—C process.} \quad (10)$$

$H(B \to C)$ is variety of B-C combinations. Small $H(B \to C)$ may means B-C process is highly guided.

$$T(B \to C) \equiv H(B) + H(C) - H(B \to C) \quad \text{: Amount of information transfer.} \quad (11)$$

T(B→C) is effective amount of information that the user can input to the machine correctly.

Concepts discussed so far can be found in the paper of Fitts. Fitts used only information transfer, T, to compare usability of input devices. He just thought large T means good human machine interaction. He did not care about further meaning of information transfer.

We can see practical meanings of information transfer, T, as logarithm of necessary and sufficient number of internal state of the input device for the user and the task:

$$(\text{Proper number of internal state}) = 2^{T(B \to C)}. \quad (12)$$

From a point of view of informatics, input devices with exceeded DoF are judged as redundant.

Depending types of buttons, we can transform T to get convenient indexes;

$$(\text{Proper number of buttons of interface consisting only of exclusive on/off buttons}) = 2^{T(B \to C)} \quad (13)$$

$$PN_2 = T(B \to C) \quad \text{: Necessary and sufficient number of non-exclusive, on-off-state buttons.} \quad (14)$$

$$PN_n = PN_2 \cdot \log_n 2 \quad \text{: Necessary and sufficient number of non-exclusive, n-state buttons.} \quad (15)$$

We can also derive two types of normalized T called information transfer efficiencies;

$$Dist(B \to C) \equiv \frac{T(B \to C)}{H(B)} \quad \text{: Distinguishablility of B} \quad (16)$$

Distinguishablility means ratio that difference of B influences on C selection.

$$Cont(B \to C) \equiv \frac{T(B \to C)}{H(C)} \quad \text{: Controllability of C} \quad (17)$$

Controllability means ratio that occurrences of certain C is controllable by selecting proper B.

In flawless human inputs, information transfer efficiencies become one. Erroneous or unguided behaviour of users makes information transfer efficiencies almost zero.

Table 1. User behavior. Frequency of pair of buttons and levers commanded to operated and actually pressed

		Operated Button/Lever																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Button/Lever commanded to operate	1	30	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	2	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3	0	0	26	6	1	0	0	0	0	5	2	0	0	0	0	0	0	0	0
	4	1	0	0	31	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	1	36	0	0	0	0	0	2	0	0	0	0	0	0	0	0
	6	0	0	0	1	0	33	0	0	0	3	0	0	0	0	0	0	0	0	0
	7	1	0	1	0	0	0	31	0	0	0	1	0	0	0	0	0	0	0	0
	8	0	0	0	0	0	0	1	21	0	0	0	0	0	0	0	0	0	0	0
	9	0	0	0	0	0	1	0	0	30	0	0	0	0	0	0	0	0	0	0
	10	0	0	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0
	11	0	0	0	0	2	0	0	1	0	0	36	0	0	0	0	0	0	0	0
	12	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0
	13	0	0	0	0	0	0	0	0	0	0	0	0	29	0	0	0	0	0	0
	14	0	0	0	1	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0
	15	1	0	0	0	0	0	0	0	0	1	0	0	0	0	32	0	0	0	0
	16	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	27	0	0	0
	17	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0
	18	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	29	0
	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	31

3 Example of evaluation

3.1 Input device to be evaluated

I employed a game controller, TAITO TCPP-20014, which is designed for tram control simulation games. The controller has eleven on-off state buttons, a five-notched lever, and an analogous-input lever (Fig. 1). Since we deals all of the button as exclusive button (radio button), the number of internal state is 19.

3.2 Procedure of experiment

Each of human subjects experiences inputting operations with the input device. A computer screen is set in front of the subject to show pictures that indicate buttons and levers to be operated. The pictures shown in the screen are obvious and easy to understand. Even though the pictures are stimuli (classified as A), they are obvious and can be connected directly to decisions of B. So we can deal the pictures as B.

The subject is asked to operate the input device for 120 seconds without seeing it. I employed the blind condition to make the operations rather erroneous. The system randomly selects buttons and levers to be operated, and the selection probabilities are uniform for every buttons and levers.

3.3 Result and Discussion

The subject group was eight novices consisting of two females and six males in age of 20—40.

Table 1 shows result of observed number of B—C pairs. By processing the data, we got analysis result shown in Table 2.

Values in the column “2^X” indicate number of on-off buttons.

Value of 2^H(B) was 18.7, i.e. the subjects were required cognitive effort to select one out of about nineteen. It agrees with the fact that the task selection was uniformly-random selecting one out of nineteen.

Table 2. Analysis result. Information transfer and number of DoF used effectively.

	Subject ID								All Subject	
	1	2	3	4	5	6	7	8	Sum or Ave	2^x
N	61	73	85	65	94	77	76	89	620	
Correct #	54	71	80	60	81	74	69	88	577	
Error #	7	2	5	5	13	3	7	1	43	
Correct Rate	89%	97%	94%	92%	86%	96%	91%	99%	93%	
H(B,C)	4.39	4.23	4.32	4.25	4.56	4.08	4.34	4.14	4.61	24.4
H(B)	4.02	4.14	4.14	3.94	4.06	3.90	4.01	4.10	4.22	18.7
H(C)	4.12	4.12	4.13	4.02	4.08	3.91	4.09	4.09	4.23	18.8
T	3.75	4.02	3.94	3.71	3.58	3.73	3.76	4.05	3.85	14.4
Dist	91%	98%	96%	92%	88%	96%	92%	99%	91%	
Cont	93%	97%	95%	94%	88%	96%	94%	99%	91%	

Result of 2^T was 14.4: the subjects could make task that is equivalent to select one out of about fifteen. The subjects lost information and failed to use 4.3 (= 18.7 – 14.4) buttons and levers. In other word, about four buttons are too much for human cognitive ability, and proper number of input device DoF is fifteen in this case.

The information transfer efficiencies were 91%. The subjects must make ten operations to input nine commands.

We can also analysis each error separately. Non-zero cells out of diagonal elements in Table 1 indicate operational errors. By assuming each of them as zero, we can estimate imaginary amounts of information transfer if the errors did not occur. For instance, there were six slips mistaking button No.4 as No.3. By assuming the slips did not occur, we get T = 3.88. The 2^T is 14.7, and it is 0.3 larger than actual value of 14.4. It means that the designer can increase usability and add 0.3 buttons to the device, when the slips are completely prevented.

4 Conclusion

This paper proposed a method to count number of buttons that are effectively used. The method is quantitative and universal to evaluate usability of various input devices. An example of application shows how many buttons the users lost in the operation.

References

Fitts, P. W. and Seeger, C. M. (1953) “S-R compatibility: spatial characteristics of stimulus and response codes,” *Journal of Experimental Psychology, 46,* 199—210.

Garner, W. R. and Hake, H. W. (1951) “The amount of information in absolute judgments,” *Psychological Rev., 58,* 446—459.

Huffman, D. A. “A Method for The Construction of Minimum Redundancy Codes,” *Proceedings of IRE, 40(9),* 1098—1101.

ISO (1997) “ISO 9241-14. Ergonomics — Office work with visual display terminals (VDTs) — Menu dialogues.”

Miller, G. A. (1956) “The magical number seven, plus or minus two: some limits on our capacity for processing information,” *Psychological Review, 63,* 81—97.

Norman, D.A. (1988) *The design of everyday things*; Basic Books.