Preference Aggregation Based Cognitive Modeling: An Alternative Explanation of the Wason Selection Task*

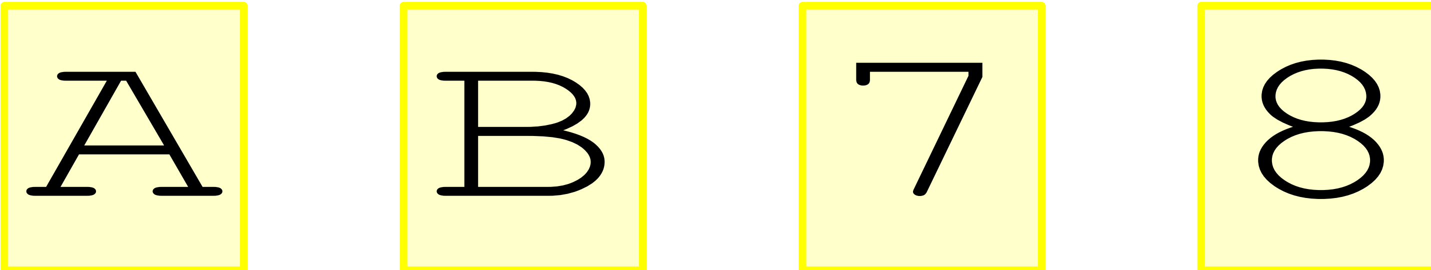
Kenryo Indo
Department of Management, Faculty of Economics, Kanto Gakuen University, Ota-si, Gunma-ken 373-8515, Japan

*) This handout was delivered at COGSCI 2008,
23-26 Sep 2008, Omni Shoreham, Washington D. C.

1. Motivation

- Wason's four-card selection task (hereafter the WST) shows that it can be very difficult for human reasoner to validate (correctly, falsify) a simple conditional sentence.
- This problem has been attracted many researchers. For example, Mental Model (Johnson-Laird and Byrne, 2002), Pragmatic Reasoning Schema (Cheng and Holyoak, 1985), Relevance Theory (Sperber et al., 1995), Dawinian Algorithms (Fiddick et al., 2000), Dual Process Theory (Evans and Over, 2004).
- I propose an alternative explanation of WST based on preference aggregation, which has been studied by social choice theorists (Arrow, 1963; Sen, 1982; Gaertner, 2001), with a PROLOG implementation.

The Wason Selection Task (WST)



Rule: “If there is an A on one side of the card, then there is a 7 on the other side of the card.” Choose those cards that need to be turned over to decide whether the rule is true or false.

Fig. 1. A version of the Wason selection task.

- The unique violation is the case of A and 8 (“p” and “not q” in the schematic form)
- Correct solution rate is typically 10% or less (ex., Evans and Over (2004); Evans and Lynch (1973)).
- Most of participants select only A, or both A and 7 (in schematic form, “p”, or “p & q”).
- Human intelligence tends to be affected by context or contents of the problem.

2. Modeling the WST as a decision making problem

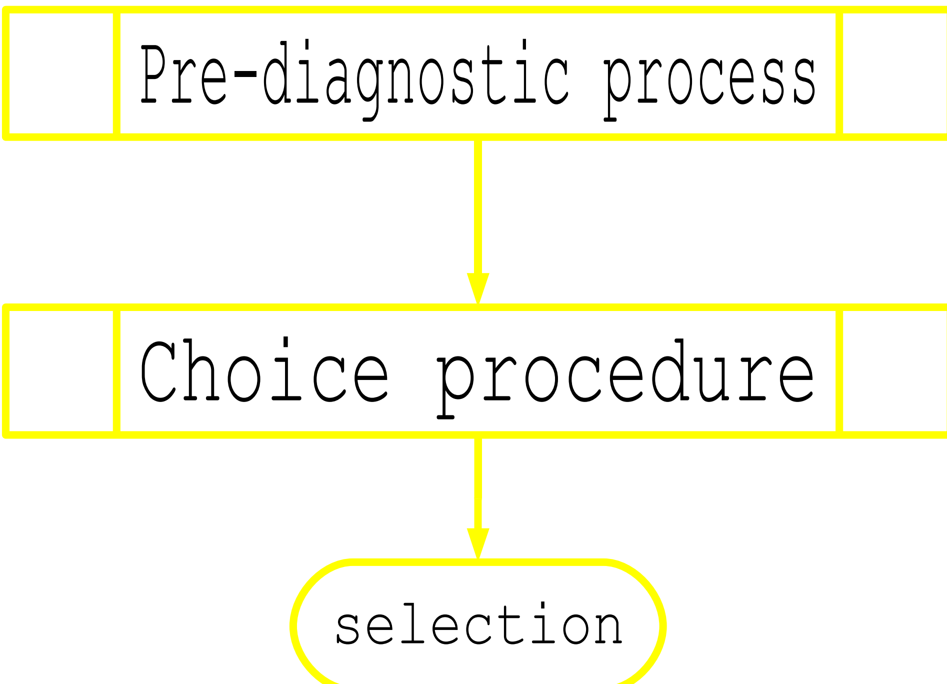
- In our modeling, firstly, WST is translated into a two-stage decision making problem.
- The subject of WST can be seen as a decision maker (DM) who should decide whether to select or not for each card (it can be represented by the following function d).

- d:D⟶X

- d is a function of the data set D to the action set X, where X={inspect, not inspect}, D={p, q, not p, not q}

pre-diagnostic process

- Other than the truth table, some **procedural knowledge** seems to be needed in solving the WST. I would like to call this setup stage the ***pre-diagnostic process***, followed by the subsequent *choice procedure*.



Definition

- We assume that before each inspection, a DM constructs the knowledge representation (or the rights system) which provides information to solve the task.
- Formally, a *pre-diagnostic process* is a function, h:D$\cup$\{T\}$\rightarrow$M, where D is the data set, T is a reservation level of inspection, and M is interpreted as *knowledge representation* of the DM. For any data d, we call h(d) a *concerning set*, or a rights system of the DM. The *choice procedure* is defined as g:M$\rightarrow$X, so d=g·h.

Table 1. aggregation of the selection task

Reason (voter)	p	q	$p \to q$	Reason (voter)	not p	not q	$q \to p$
R1	1	1	1	R1	0	0	0
R2	1	0	0	R2	0	1	1
R3	0	0	1	R3	1	1	0
M (majority rule)	1	0	1	M (majority rule)	0	1	0

1: inspect the card

0: do not inspect the card

- the probability of cycle is about 5.6% for 3-person and 3-candidate and less than 8.8% for any 3-candidate cases (See Gaertner (2001) p.37, Table 3.1).

3. Aggregation

- I will model the pre-diagnostic process *h*, and the succeeding choice procedure *g*, for WST as a preference (i.e., ordering) aggregation.
- In Table 1, unit values and zero values represent *"more suspicious than" relations over card data and a threshold value (or the truth)* and *default beliefs* respectively. These are three individually alleged reasons by R1, R2, and R3, and M the aggregated ordering according to pairwise majority, observing a data in each column label (See Table 1).
- Table 1 may be seen as 'doctrinal paradox' (Kornhauser, 1992), or 'judgment aggregation' (Diederich, 2006), which is intensively studied recently. But here I regard this not as a paradox but rather a cognitive modeling of (mis)understandings.

Cognitive rights system: An interpretation of orderings

- In our modeling, it is assumed that the *pre-diagnostic process* generates a knowledge representation system in order to distribute the *justifications of inspection* (1-values), and the *protected privacies* against inspection (0-values), with respect to "*more suspicious than*" relations, respectively in the majority row of Table 1.
- Note that the orderings in this paper are not merely preference relation used in economics rather interpreted as the *rights*, or *relative relevance for inspecting cards*. It should not be executed if it harms other important rights.

4. Computational model

- Each PROLOG program in Figure **4** respectively stands for , 0-valued propositions (on-belief) of R1, of R2, and of R3, in Table 1. And their failed queries represent 1-valued propositions (off-belief) to be inspected for the cards in Table 1.

```
%R1
r(1):p :- r(1):q.
r(1):not_p.
r(1):not_q.
```

```
%R2
r(2):q :- r(2):p.
r(2):q.
r(2):not_p.
```

```
%R3
r(3):p :- r(3):q.
r(3):p.
r(3):q.
```

```
?- to_inspect.
1:  [p][q]
2:  [p][not_q]
3:  [not_p][not_q]
M:  [p][not_q]
```

Fig. 4. PROLOG implementation of the three orderings, which are represented as Table 1. And the rights to inspect proved by resolution. See Appendix for the source code

5. Simulating the Biases

- The following r(4) a slightly modified version of r(3), which stands for R3 in Table 1, together with r(1) and r(2), simulates the *confirmation bias* (or the matching bias in the case of affirmative indicative conditional) as a result of majority decision (denoted as M(i-j-k)).

```
r(4):p :- r(4):q.
r(4):p.
r(4):not_q.
--> 4: [q][not_p]
M(1-2-4): [p][q]
```

• R4 is just reversing the values both for q and for not_q in R3, and R5 is just reversing the values for not_p and for p in R2, respectively, in Table 1.

• Similarly, each of the two orderings r(5) and r(6) is the revision of r(2) and of r(1) respectively.

```
r(5):q :- r(5):p.
r(5):p.
r(5):q.
--> 5&M(1-5-3): [not_p][not_q]
```

```
r(6):q :- r(6):p.
r(6):not_p.
r(6):not_q.
--> 6: [p][q]
M(6-2-3): [not(q)][p]
```

- However, it **is impossible** to produce **p as the single winner** by a majority, **except for a pair of R1 and R2**, and its variants.

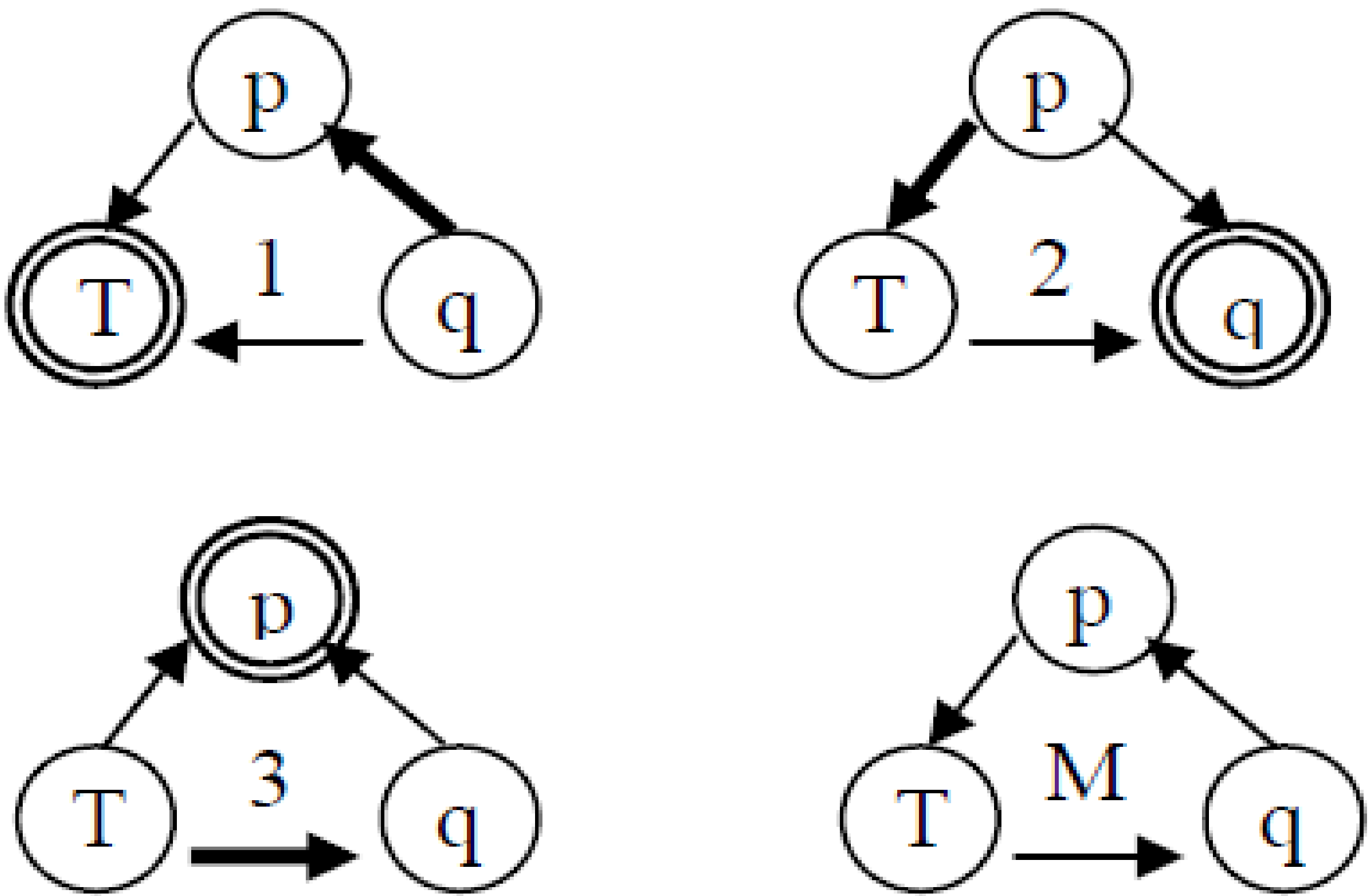


Fig. 3 a) Another graphical view of orderings in Table 1.

- r(1) : $T\ (\leftarrow \boxed{F) \leftarrow p \leftarrow q}$
- r(2) : $q \leftarrow \boxed{T\ (\leftarrow F) \leftarrow p}$
- r(3) : $p \leftarrow \boxed{q \leftarrow T\ (\leftarrow F)}$
- r(4) : $p \leftarrow \boxed{T\ (\leftarrow F) \leftarrow q}$
- r(5) : $q \leftarrow \boxed{p \leftarrow T\ (\leftarrow F)}$
- r(6) : $T\ (\leftarrow \boxed{F) \leftarrow q \leftarrow p}$

Fig. 3 b) Ordering malleability

Propositions after T are not to be inspected. Each malleable ordering may reverse a unique direction without changing the top object.

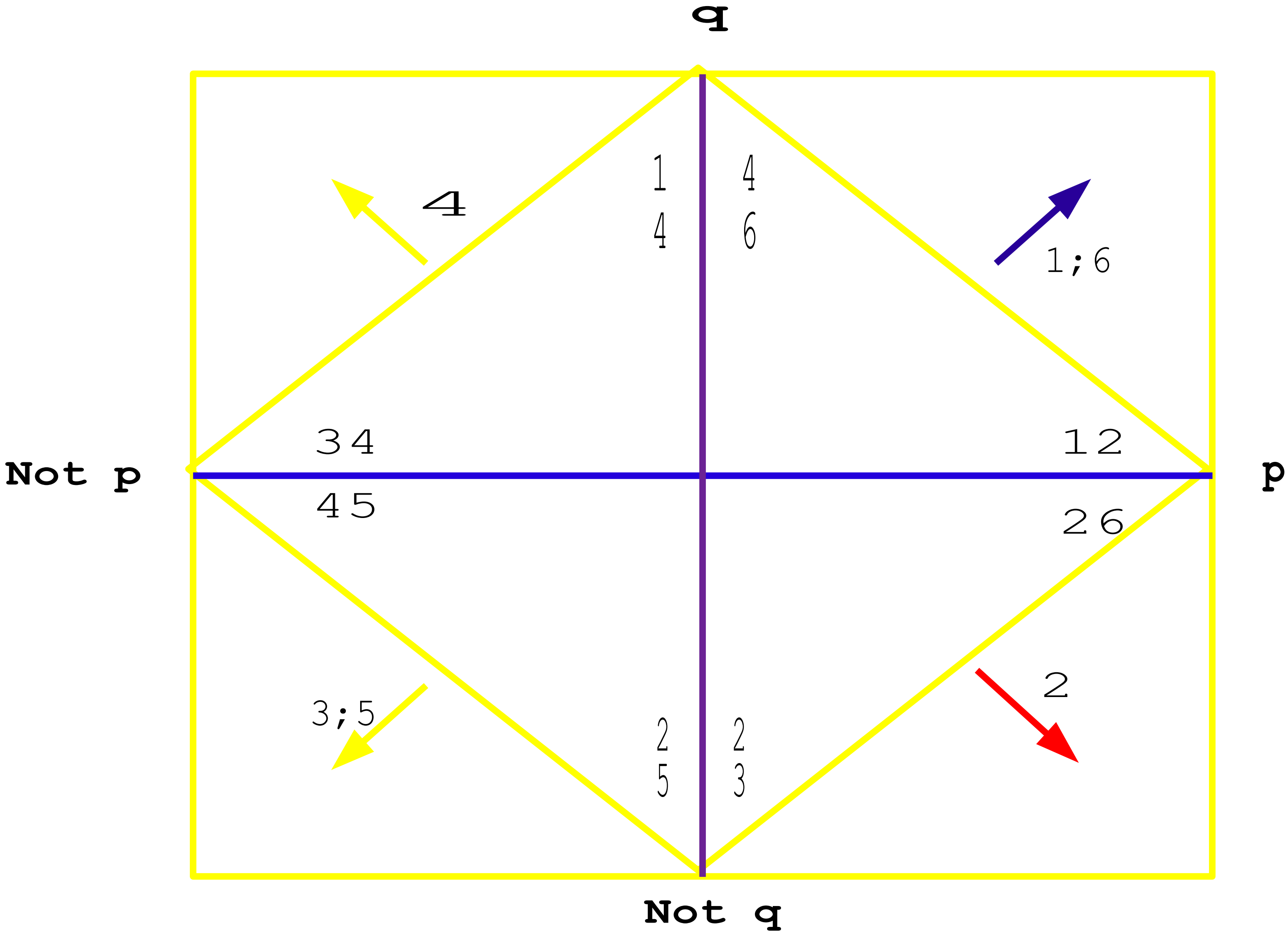


Fig. 3 c) Map of the majority winner (i.e., the most suspicious one) for the WST

Table 2. Simulating the selection probabilities in Evans and Lynch (1973) for if p then q.

ordering before decay	TA	TC	FA	FC	SUM	cases	TA	TC	FA	FC
1256	66	38	12	38	154	80	83%	48%	15%	48%
1236	66	38	12	38	154	80	83%	48%	15%	48%
126	39	26	0	9	74	39	100%	67%	0%	23%
26	14	6	0	6	26	14	100%	43%	0%	43%
12	14	6	0	6	26	14	100%	43%	0%	43%
total	199	114	24	97	434	227	**88%**	**50%**	**11%**	**43%**

Evans and Lynch (1973)

	TA	TC	FA	FC	SUM	cases	TA	TC	FA	FC
If p then q	21	12	2	8	–	24	88%	50%	8%	33%
If p then not q	22	2	1	14	–	24	92%	8%	4%	58%
If not p then q	14	14	7	10	–	24	58%	58%	29%	42%
If not p then not q	13	7	11	18	–	24	54%	29%	46%	75%
Overall%	70	35	21	50	–	96	73%	36%	22%	52%

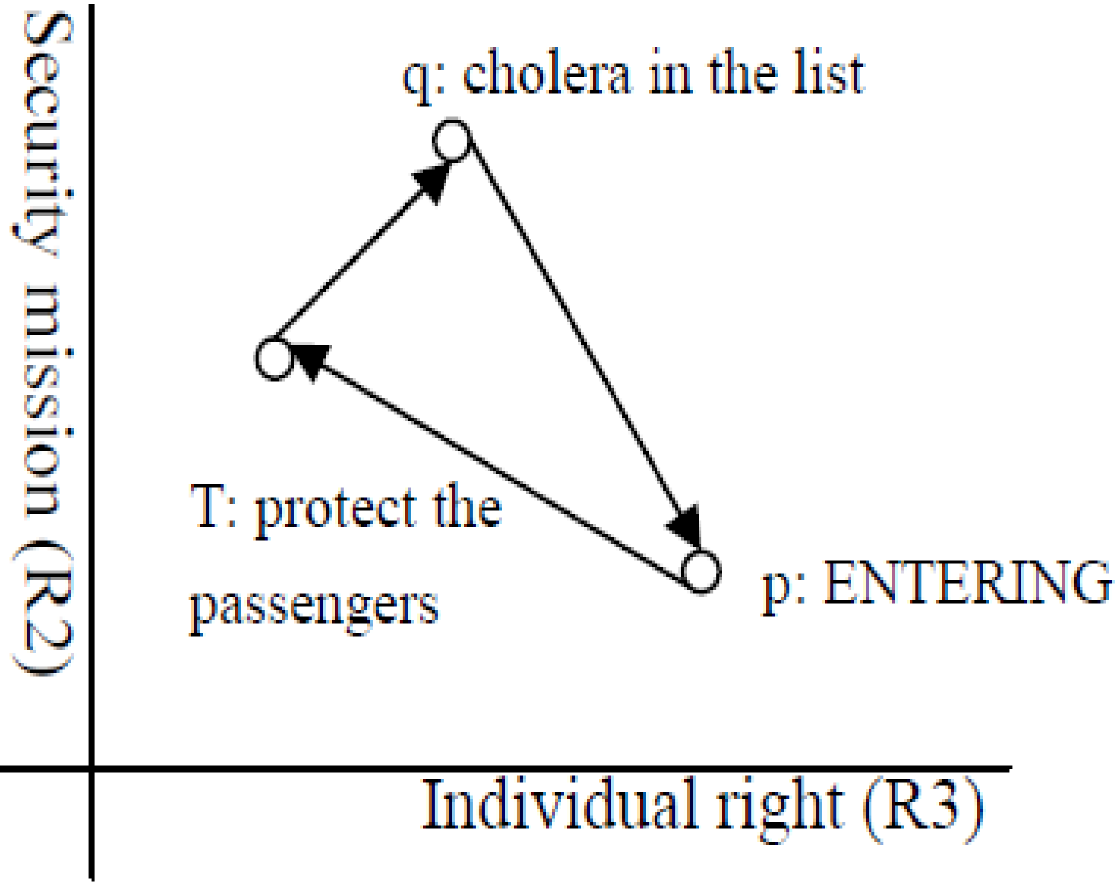
TA, FA:
True Antecedent
False Antecedent

TC, FC:
True Consequence
False Consequence

6. Cognitive stability

- Ordering Malleability. The three orderings in Table 1 consists a Latin Square. Reversal of a direction for each arrow transforms it into another (linear) ordering (See Figure 3). In Figure 3, three bold arrows can be reversed locally without changing their top-level rights (doubly circled), respectively. We will say these bold arrows (and the ordering) are *malleable* with respect to the cognitive rights system.
- As shown by social choice theorists, any non-manipulable and non-dictatorial pairwise choice defined on the restricted domain (a subset of orderings) is also (a part of) pairwise majority vote. And this result can be generalized to models larger than 2-agent and 3-alternatives. So, it is *stable* against malleability of a single relation unless it touches the Latin Squares. (* Gibbard-Satterthwaite Theorem, and Maskin-Campbell-Kelly Theorem)

- In summary, the WST is considered to be *self-deceptive* in that the subject selects the expected answer only if he/she sticks to an instable, cyclic majority decision.
- Lastly, we turn our attention to the pragmatic reasoning schema, especially the permission schema by Chen and Holyoak (1985).
- Passengers at an airport were required to show a form with *a list of diseases*, and it is necessary to check whether the following rule is violated.
  - “If the form says ‘ENTERING’ then ‘*cholera*’ is included in the list.”
- It is shown that the performance was sensitive to whether with or without suggesting the rationale “to protect the passengers against the disease”.
- A part of the Latin Square in Table 1 is considered to be naturally consisted under the above rationale. See Figure 4.



Geographically this ordering profile is similar to the asymmetric dominance (Simonson, 1992) in the consumer choice theory. It is also a critical profile of the Paretian Liberal by Sen (1982).

Fig. 4 The permission schema as coordinating the conflicting rights.

The three rights just consist the Latin square in Table 1, if we suppose *additionally R1 which is not appeared in this figure, as a coordinator, balancing two conflicting rights.*

References

Arrow, K. J. (1951/1963). Social Choice and Individual Values, Second Edition, Yale University Press.

Campbell, D. E. and Kelly, J. S., “A strategy-proofness characterization of majority rule,” Economic Theory, Vol. 22, pp. 557-568, 2003.

Cheng, P. W., and K. J. Holyoak (1985). Pragmatic reasoning schemas, Cognitive Psychology, Vol.17, pp. 319-416.

Clocksin, W. F. and C. S. Mellish (2003). Programming in Prolog: Using the ISO Standard, 5th edition, Springer.

Dietrich, F. (2006). “Judgment aggregation: (im)possibility theorems,” Journal of Economic Theory Vol.126(1): 286-298.

Evans, J. St. B. T. (1989). Bias in Human Reasoning: Causes and Consequences, Lawrence Erlbaum.

Evans, J. St. B. T., and J. S. Lynch (1973). Matching bias in the selection task, British Journal of Psychology. 1973 Aug Vol Vol. 64(3), pp. 391-397.

Evans, J. St. B. T., and D. E. Over (2004). If, Oxford University Press.

Fiddick, L., L. Cosmides, and J. Tooby (2000). No interpretation without representation: the role of domain-specific representations and inferences in the Wason selection task, Cognition, Vol.77, pp. 1-79.

Gaertner, W. (2001). Domain Conditions in Social Choice theory, Cambridge University Press.

Holland, J. H., K. F. Holyoak, R. E. Nisbett, and P. R. Thagard (1986). Induction: Process of Inference, Learning, and Discovery, MIT Press.

Johnson-Laird, P. N. and R. M. J. Byrne (2002). Conditionals: a theory of meaning, pragmatics and inference, Psychological Review, Vol. 109, pp. 646-678.

Kornhauser, L. A. (1992). Modelling collegial courts II: legal doctrine, Journal of Law, Economics and Philosophy, Vol. 8, pp.441-470.

Popper, K. (1959). The Logic of Scientific Discovery, Hutchinson.

Sen, A. (1982). Choice, Welfare and Measurement, MIT Press.

Simonson, I. (1989). Choice based on reasons: the case of attraction and compromise effects, Journal of Consumer Research, Vol. 15, pp. 158-173.

Sperber, D., F. Cara, and V. Girotto (1995). Relevance theory explains the selection task, Cognition, Vol.57, pp. 31-95.

Taylor, A. D. (2005). Social Choice and the Mathematics of Manipulation, Cambridge University Press.

Wason, P. C. (1966). Reasoning, In B. M. Foss (ed.), New Horizons in Psychology I, pp. 135-151, Penguin.