



## WHY IS RASE A FOUR LETTER WORD? SEEKING A THEORETICAL BASIS FOR A CONCEPTUALISATION OF THE KNOWLEDGE FOUND IN DOCUMENTS

Nicholas Nisbet and Selçuk Çidik  
University College London, London, UK.

### Abstract

The RASE (requirements, applicability, selection and exception) approach has been developed as a means to make regulatory and other text machine-operable whilst remaining human-readable. RASE manifests as coloured mark-up of sections and phrases overlaid on the existing text and tables, making normative documents accessible for systematic checking. However it has been perceived as lacking theoretical basis or justification. In particular the question can be asked as to why there are four RASE colours needed, and not a larger or smaller number. This simple question raises important questions concerning both research methodologies and about learning processes.

### Introduction

Work in the field of automated compliance checking had proceeded without RASE and without even foreshadowing RASE, by depending on the authoring of procedural code, or the filling of predefined templates. The conventional use of mark-up has been focussed on the exploitation of HTML for presentation on the world wide web with the addition of cross-referencing. Since its initial publication (Nisbet et al, 2008), RASE has been subject to further academic investigation. Hjelseth et al developed a series of papers (Hjelseth et al 2010, 2011) which developed the initial publication. More recently Nisbet et al have published a series of papers which have explored some implications, starting with an overview of the logic of knowledge embedded in normative documents (Nisbet et al, 2019), the use of RASE document to guide design and the handling of uncertainty (Nisbet et al, 2021), the mapping of RASE document to other formal representations and languages (Nisbet et al, 2022a), the direct execution of RASE documents against models (Nisbet et al, 2022b) and the use of RASE to capture and execute semantic enhancement of models (Nisbet et al, 2024).

There is now also a growing body of work where RASE has been successfully applied since its first use in the ICC 'SmartCodes' project (Nisbet et al, 2008) where the aim was to eliminate the mutual dependence between regulatory experts, programming experts and model

experts when creating automated compliance checking (ACC) systems. In some cases RASE has been used as a discipline to structure thinking about regulations prior to manual transcribing specific sentences into procedural code. In the regBIM (Beach et al, 2020), DCOM (Beach et al, 2019) and ACCORD (Hettiarachchi et al, 2024) research and development projects the RASE approach has been used to embed the analysis into the document and automatically generate procedural, semi-procedural or declarative content for consumption by rule engines. In the ACCORD project, some progress has been made in automating the application of RASE to sentences using Machine Learning (Al-Turki et al, 2024).

RASE is an approach to the analysis to text and tables that seeks to identify the individual metrics in phrases and words and the distinct objectives in sections found in normative regulations, and also in definitive resources and informative documents. RASE does not require that a document be purely normative, definitive or informative. Depending on the editorial rigour with which the document was developed, all or some of these three types of knowledge may be present.

The metrics and objectives found in a document are marked-up as Applicability, Selection, Exception and either as Requirements, Registrations or Reports, depending on the modality of the content. The RASE mark-up can be given distinct colours to support review and correction of otherwise monochrome documents (see Figure 3 below). The document containing RASE mark-up can be consumed directly by applications or can be transformed into other logical or procedural representations that can be consumed by rule-engines or compilers.

The execution of RASE for compliance checking in the built environment is based on the insight that any object, from a site or facility down to the smallest constituent parts may satisfy any regulation, from a whole document down to an sentence or an individual cell in a table, in four distinct ways. The most obvious is that the object is 'as Required' by the regulation. This might be noted as a tick on a checklist form. There are an additional three ways, namely 'not Applicable', 'not Selected' or 'Excepted'. These three may be noted by the catchall 'n/a' on a

checklist form, but they are actually distinct, as shown in Figure 5. The converse statement can be made, that for an object to not satisfy a regulation it must be ‘Applicable’, ‘Selected’, ‘not Excepted’ and ‘not as Required’.

Each of these four aspects can be considered in detail as each of the four types of content has its own effects when interpreted. There may be none, one, or several phrases or clauses that define the Applicability. If there are several Applicabilities then any one may be failed to achieve ‘not Applicable’. There may be none or several phrases or clauses that define the Selection, but not just one. If there are Selections then all must be failed to achieve ‘not Selected’. There may be none, one or several phrases or clauses that introduce Exceptions. If there are several Exceptions, then any one may be met to achieve ‘Excepted’. There will always be at least one but possibly several phrases or clauses that define normative Requirements, definitive Registrations or informative Reports. If there are several, then each is expected to be met to achieve ‘as Required’.

### **Problem statement**

When presenting or teaching the RASE approach to students or subject matter experts, there is frequently an initial stumbling block in the proposition that exactly four ‘colours’ are needed. This seems in contrast to the common knowledge that text in general and regulations in particular are untidy and intractable. Those students that have had some experience of grammatical analysis, either for parts-of-speech or for sentence structure, are particularly sceptical as these analyses are open-ended, there being no limit to the number of parts-of-speech nor to the complexity of structure. The question about ‘four-ness’ is therefore of importance in its own right and also as a question that leads to how a novel conception of knowledge in documents can be justified when it has no direct precedents. Without precedents, many of the conventional modes of academic research are challenged. This paper will seek to address the iceberg question of ‘four-ness’ – why are there four colours needed representing four kinds of content? - as a gateway, since this may support the future exploration of the larger questions.

### **Approach**

This is necessarily an informal and ‘preliminary’ investigation into what might constitute a validation for a novel conceptualization of knowledge in documents, which is not derivative from past work, nor an induction from a body of past observations. The simple question “Why four?” is the most commonly heard expression of the anxiety on confronting the novelty and broadness of the RASE approach, but it opens the question as to what would count as an appropriate methodology and what might constitute a good answer to both this specific question and to the broader issues raised by what could potentially be a shift in how we think about knowledge in

normative, definitive and informative content in documents.

### **Analysis**

Humans can accept apparent ‘patterns’ unthinkingly. RASE has attracted a number of such simplifications. We can consider the ‘four-ness’ of RASE in a number of ways. Some of these may lack scientific rigour, or, whilst apparently scientific, may on examination prove less well founded.

### **Using numerology**

It is important to guard against finding justification or assurance in numerology. Historically there are examples of ‘four-ness’ as a given that now cannot be revised due to their location in the past, such as the ‘humors’ or ‘elements’ in ancient Greek medicine. Examples of the numerology around ‘four-ness’ include its use in various schemes to represent stability. Pythagorean mysticism associated ‘four-ness’ to harmony, proportion and stability. Another example would be ‘quaternity’ in Carl Jung’s constructed mythology as representing wholeness and completeness but also fairness and equality. In Feng Shui various balances of four directions, four celestial animals and four pillars of destiny are sought. Any such association can be countered by reference to Chinese mysticism where ‘four-ness’ is considered unlucky because its phoneme is similar to that of death. In other contexts, that there were seven colours in the rainbow and seven dynamic heavenly bodies are examples where research has been truncated because of numerological capture. Being aware of the subtle influence of numerology can remind us to not stop searching for more options such as ‘five-ness’ nor not stop pruning back to ‘three-ness’ just because we have arrived at what might be an attractive number of options.

### **As fact**

RASE can be presented and taught as a ‘de facto’ statement based on the pragmatics of its past successful usage. Users of the methodology have accepted the ‘four-ness’ as a working or given assumption. This could suggest that the ‘four-ness’ question is no more valid when applied to RASE than it would be to points-of-the-compass or the number of fundamental forces, currently true but not proven, and so potentially subject to revision. This conclusion is unsatisfactory if there is to be investment in the ideas and implementation of RASE.

### **Using analogy**

Hjeljseth (2010) has made an analogy between the four colours in RASE with the Aristotelian ‘Square of Oppositions’. This identifies four kinds of universal statements, whilst RASE identifies four kinds of specific phrases. These can be linked together with six possible connections as with the points and edges of a tetrahedron.

In both cases the tetrahedron can be projected onto the page as a square, as shown in Figure 1.

A <b>Universal Affirmative</b> All S are P	E <b>Universal Negative</b> No S are P	R (blue) <b>Requirement</b> Shall be / Is defined as / Is reported as R's	E (orange) <b>Exception</b> Not these Es
I <b>Particular Affirmative</b> Some S are P	O <b>Particular Negative</b> Some S are not P	S (purple) <b>Selection</b> Some of these Ss	A (green) <b>Applicability</b> These As

**Relationships:**

Diagonals: AO and EI	Must be opposite	Diagonal: RA	Union (AND)
Top row: AE	Can both be false	Diagonal: SE	Intersection (OR)
Bottom row: IO	Can both be true	Top row: RE	Object (mostly) (-)
Columns: AI and EO	Truth sinks False floats	Bottom row: SA	Subject (NOT)

Figure 1: Comparisons of Aristotle and RASE squares

Such squares can be arranged in several different ways so that making analogies can be arbitrary. Analogy can act as a reassurance that a particular suggestion is plausible in mimicking some other pattern, or can boost creativity by creating a challenge to find the similarities and differences. Here the similarities are that four topics, or indeed any four topics, can be linked with six possible links. The differences are that one deals in abstract concepts, the other written text and tables. Aristotle addresses the universal statements in informative content whereas RASE addresses specific normative and definitive as well as informative content. When considering the six relationships, there are few direct similarities. There is the risk of over reliance on the analogy as the differences may be less well considered.

**Using grammar**

Since RASE is a structure identified in written text and tables, it may be possible to rationalise the ‘four-ness’ by reference to grammatical roles of phrases and words in sentences. An example might be the subjects and objects of conventional grammatical parsing into subject, verb and object. In simple sentences using the active sense in English, Mandarin and most other languages, the initial phrases or words identify the subject(s), whereas phrases and words after the verb identify the object(s). If each of these two constructs were to have two variants then we would arrive at ‘four-ness’. In the case of RASE, subjects may be narrowed with increasing specificity of Applicability, or may be widened with the identification of a Selection of different subjects. The object of such

sentences may be a Requirement in normative sentences, a Registration in definitive sentences or a Reporting in informative sentences. Alternatively the object of a sentence may be modified by an Exception. However exceptions can also apply to the subject scope. So although we have identified four types, the ability of exceptions to affect either the subject or the object means we have not detected a justification of ‘four-ness’.

**Using pedagogy**

One of the drivers for this investigation is to support the teaching and learning of RASE. To date there have been no structured investigations into the pedagogy, i.e. teaching and learning, of RASE to human audiences, though informally it needs at least thirty minutes to engender an initial confidence and longer to achieve proficiency. We can conduct experiments by simulating the human teaching process by targeting a LLM. This can involve first focussing the LLM by requesting that it perform conventional parts-of-speech analysis of the chosen text followed by performing sentence-structure analysis. This refreshing of language terminology ensures that the following prompt (Figure 2) is correctly interpreted on a synthetic example (Figure 3).

Add mark-up to any adjective, noun, proper noun or verb participle phrase or word including any preceding preposition, when appearing anywhere after 'shall' with 'Requirement', With 'Exception' when appearing anywhere in a phrase after 'except', with 'Selection' when appearing anywhere in any phrase containing a list that includes 'and' or 'or', and with 'Application' when appearing anywhere else.

Figure 2: LLM prompt

Every blue rectangular wooden door , window or trap unless in London , protected or temporary shall be accessible , automatically-operated and self-closing .

Figure 3: LLM response

This prompt can correctly mark-up some chosen examples but is only a partial solution, with no attempt to deal with more than one sentence at a time, and with only a few specific conjunctions and verb constructs covered. Further refinement of the prompt may improve the generality of the approach and provide further guidance for how the teaching of RASE could be conducted.

**Using logical analysis**

Rather than seeking explanation of the ‘four-ness’ of RASE in grammatical analysis, it may be useful to conduct a study of the logical evaluation. Most traditional works and textbooks use AND, OR and NOT, which are broadly analogous with words used in English and many other languages, though a few languages achieve these meanings by using tone and context, or by using repetition of whole sentences with different subjects or objects. NOT is a unary ‘prefix’ operator, occurring before its term. AND and OR are binary ‘infix’ operators occurring between two terms. Figure 4 illustrates this usage.

Pass objects that are NOT a door OR are waterproof

Figure 4: Example logic operators in natural language

The need to consider lists of more than two terms (especially in electronics) has led to explorations of the four set logical operators ‘AND, OR, NOTAND and NOTOR’. These can be used as functions using prefix notation. These have a correspondence to the way the four RASE ‘colours’ operate on each individual phrase or subsection (Figure 5, equations 1a to 1r):

Result =  
 OR (‘not *Applicable*’, ‘not *Selected*’, ‘*Excepted*’, ‘as *Required*’) (1)

‘not *Applicable*’ = NOTAND ( $A_1, \dots A_n$ ) (1a)

‘not *Selected*’ = NOTOR ( $S_1, \dots S_n$ ) (1s)

‘*Excepted*’ = OR ( $E_1, \dots E_n$ ) (1e)

‘as *Required*’ = AND ( $R_1, \dots R_n$ ) (1r)

Figure 5: Five RASE logic equations

These four operators also have the important feature of being both associative and commutative. Commutativity means that the values can be considered in any order, so we can respond as knowledge becomes available or adopt heuristics to test the more decisive metrics earlier. Associativity means that given the evaluation of any proceeding terms, we can evaluate the next directly. Examples are shown in Figure 6.

Commutativity:	‘waterproof and fireproof’ is the same as ‘fireproof and waterproof’.
Associativity:	‘waterproof and fireproof and also windproof’ is the same as ‘waterproof and also fireproof and windproof’

Figure 6: Examples of commutativity and associativity

However, by admitting these four logical operators as candidates, we should also consider other logical operators. In the extreme case, it is possible to define  $2^x$  logical operators when considering x values. Each of these are valid logical operators that can be applied to lists of values, such as the results of evaluating individual metrics or objectives in sentences. Most of these arbitrary (and uncountable) operators can be dismissed because the majority do not treat the individual values equally, whereas a criteria for a systematic approach to knowledge is to not privilege any individual metric over others. There are four further operators that reflect this expectation which have been named, namely XOR (exclusive or), XNOR (exclusive not or), MAJority and MINority. XOR is usually used in binary situations to express that, given two alternatives, either is exclusive to the other. XOR has also been extended to include the idea of oddness in contrast to evenness over any list of values. XNOR indicates if all the values are the same. MAJority indicates

if more than half of the values are true and conversely MINority indicates if less than half of the values are true. Other operators could be defined for ‘super-majority’ voting such as over 2/3rds, as found in some constitutions. These additional candidate operators have a common feature not present in the four operators currently used by RASE. These logical extensions require oversight in their evaluation, in the form of a separate agent, often called a ‘teller’ with the duty to count votes. Since we expect knowledge to be expressed in terms independent of an outside actor, we can exclude them as lacking ‘decidability’. We can therefore propose that the four ‘colours’ in RASE do not need to be expanded further because any additional colour and associated logical operator would not meet the expectation for commutativity, associativity and decidability.

## Discussion

Taken together the grammatical analysis and this detailed logical analysis allows us to say that there are only four colours found in normative, definitive and informative knowledge. However we have identified a specific exclusion. RASE may not be sufficient represent in situations where arbitrary agents are involved (breaking the expectation for decidability) or where the knowledge changes whilst being evaluated (breaking the expectation for associativity and commutativity). Examples might include constitutions designed to enforce multiple ideas of fairness, or in negotiations where facts may be promoted or relegated. Another example could be situations where votes are cast and then counted, or where votes are cast progressively. Hence we can say that RASE has a strong role in representing knowledge and in preparing the facts for judicial or political review processes but not necessarily in representing the review processes itself.

## Example implementations

In considering automatic and supported code compliance checking, we can use RASE at four distinct steps of a building permitting process based on a submission (see Table 1).

Table 1: Decision points in compliance checking

Step	Decision
1:	Is a submission needed?
2:	Is the submission valid?
3:	Is the submission compliant?
4:	Is the submission approved?

Firstly, RASE can support the process of identifying if an application is needed, secondly to test if an application is complete, and thirdly if an application is compliant to regulations. In these cases a caution would be if a regulation demanded that a majority or a minimum number of conditions be met. An example would be if a

site was considered of ‘conservation importance’ if a certain number of different species from a given list are present. This boundary is illustrated in the regBIM project when automating the UK BREEAM environmental assessment process (Kasim et al, 2018). RASE was successfully applied to the provisions of the scheme but additional markup was used to define and total the available and awarded weighted points. We also need to examine the fourth step, of converting the assessment(s) into a decision. In some cases this step is rule driven and the deontic implications around obligations and permissions are a direct result derived from the first three steps. In others scenarios, the final decision is judicial or political, and other logics may be relevant. Different facts may weigh differently, or majority decisions from individuals may be taken. We can recognize this as when a panel or agent is not just required but empowered to make a decision, as that power includes breaking or developing its own decision making rules.

This research has begun to develop the formal definition of RASE and in particular to bound it. Several different approaches have been trialled but only one has produced a definitive answer. In the course of developing these approaches the importance of equal and consistent treatment of the logical values found in operable knowledge has been emphasised and the use of voting decisions excluded. Future work may propose a representation to include these.

## Conclusion

An answer to the question about the ‘four-ness’ in RASE has been obtained. This has reinforced the current usage and identified that RASE can be applied in most situations including some, but not all, decision-making processes. This result suggests that this boundary of RASE may be a good starting point for defining the appropriate scope for human intervention. Having a ‘human-in-the-loop’ cannot be the only solution to explainability and acceptability of rule-based systems, as the burden from robust rule-based systems processing large volumes of cases would be intolerable. However, human involvement in the final decision making can be both necessary and desirable and may need to take a numeric rather than logical approach.

## References

Al-Turki, D., Hettiarachchi, H., Gaber, M.M., Abdelsamea, M.M., Basurra, S., Iranmanesh, S., Saadany, H. and Vakaj, E., 2024. Human-in-the-Loop Learning with LLMs for Efficient RASE Tagging in Building Compliance Regulations. IEEE Access.

Beach, T.H., Hippolyte, J.-L., Rezgui, Y. Towards the adoption of automated regulatory compliance checking in the built environment (2020) *Automation in Construction*, 118, art. no. 103285. doi: <https://doi.org/10.1016/j.autcon.2020.103285>

Beach, T., 2019. D-COM: Digital COMPLIance: Digitisation of Requirements, Regulations and

Compliance Checking Processes in the Built Environment-Final Report.

Hettiarachchi, H., Dridi, A., Gaber, M.M., Parsafard, P., Bocaneala, N., Breitenfelder, K., Costa, G., Hedblom, M., Juganaru-Mathieu, M., Mecharnia, T. and Park, S., 2024. CODE-ACCORD: A Corpus of Building Regulatory Data for Rule Generation towards Automatic Compliance Checking. *arXiv preprint arXiv:2403.02231*.

Hjelseth, E. and Nisbet, N., 2010, November. Exploring semantic based model checking. In *Proceedings of the 2010 27th CIB W78 international conference (Vol. 54)*.

Hjelseth, E. and Nisbet, N., 2011, October. Capturing normative constraints by use of the semantic mark-up RASE methodology. In *Proceedings of CIB W78-W102 conference* (pp. 1-10).

Kasim, T., Li, H., Rezgui, Y. and Beach, T., 2018. Integrated framework to manage building’s sustainability efficiency, design features and building envelope. In *Collaborative Networks of Cognitive Systems: 19th IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2018, Cardiff, UK, September 17-19, 2018, Proceedings 19* (pp. 650-660). Springer International Publishing.

Nisbet N, Wix J and Conover D. 2008. The future of virtual construction and regulation checking, in Brandon, P., Kocaturk, T. (Eds), *Virtual Futures for Design, Construction and Procurement*, Blackwell, Oxfordshire. doi: 10.1002/9781444302349.ch17

Nisbet, N. 2019. “The logic of knowledge”. *Proceedings of the 2019 European Conference on Computing in Construction*. University College Dublin. pp. 480-481

Nisbet, N. 2021. “Using uncertainty to link compliance and creativity”. *ECPPM 2021 – eWork and eBusiness in Architecture, Engineering and Construction*. Taylor and Francis, London. ISBN 978-1-032-04328-9

Nisbet, N., Ma, L. and Aksenova, G., 2022a. Presentations of RASE knowledge mark-up. In *Proceedings of the European Conference on Computing in Construction* (pp. 184-190). University of Turin.

Nisbet, N. and Ma, L., 2022b. Using RASE semantic mark-up for normative, definitive and descriptive knowledge. In *ECPPM 2022-eWork and eBusiness in Architecture, Engineering and Construction 2022* (pp. 436-442). CRC Press.

Nisbet, N., Zhang, Z., Ma, L., Chen, W. and Çıdık, M.S., 2024. Semantic correction, enrichment and enhancement of social and transport infrastructure BIM models. *Advanced Engineering Informatics*, 59, p.102290.