Granularity in Economic Decision Making: An Interdisciplinary Review

Shu-Heng Chen and Ye-Rong Du

Abstract. In this article, we attempt to provide a review of the idea of granularity in economic decision making. The review will cover the perspectives from different disciplines, including psychology, cognitive science, complex science, and behavioral and experimental economics. Milestones along this road will be reviewed and discussed, such as Barry Schwartz's paradox of choice, George Miller's magic number seven, Gerd Gingerenzer's fast and frugal heuristics, and Richard Thaler's nudges. Recent findings from human-subject experiments on the effects of granularity on decision making will also be reviewed, accompanied by various learning models frequently used in agent-based computational economics, such as reinforcement learning and evolutionary computation. These reviews are purported to advance our thinking on the long-ignored granularity in economics and the subsequent implications for public policy-making, such as retirement plans. It, of course, remains to be examined whether the good use of the idea of granularity can enhance the quality of decision making.

Keywords: Granularity, Paradox of Choice, Chunks, Modularity, Heuristics, Nudges, Reinforcement Learning, Evolutionary Computation.

1 Motivation and Background

While the idea of granularity is already rooted in Lofti Zadeh's earlier work on fuzzy sets and fuzzy logic, it is his article, " Fuzzy Sets and Information Granularity," [53] that gives a formal notion of granularity. This notion serves as a foundation for the later development in computing with words and granular computing. However, Zadeh himself notices that the technical notion of information granulation employed

Shu-Heng Chen · Ye-Rong Du
AI-ECON Research Center, Department of Economics,
National Chengchi University, Taiwan
e-mail: {chen.shuheng,yerong.du}@gmail.com

W. Pedrycz and S.-M. Chen (eds.), *Granular Computing and Decision-Making*
Studies in Big Data 10, DOI: 10.1007/978-3-319-16829-6_3

in [53] is "in a *strict* and somewhat *narrower* sense..." (Ibid, p.3; Italics added). Very slightly he did also mention its *broad* sense.

> Taken in its *broad* sense, the concept of information granularity occurs under various guises in a wide variety of fields. In particular, it bears a close relation to the concept of aggregation in economics;... (Ibid, p. 3; Italics added)

Being economists or social scientists, our interest in granularity may not be so much related to its strict and narrow sense of granularity; instead, what interests us is its broad sense, and a more general notion given by Zadeh is as follows:

> There are many situations, however, in which the finiteness of the resolving power of measuring or information gathering devices cannot be dealt with through an appeal to *continuity*. In such case, the information may be said to be *granular* in the sense that the data points within in a granule have to be dealt with *as a whole rather than individually*. (Ibid, p.3; Italics added)

Zadeh's idea is novel and fundamental. He actually pointed out that the elementary unit of information processing is not a number (in the real space), neither a set of real numbers, but a symbol, a concept, a feeling, a linguistic variable, a module, or a chunk, *a whole rather than individuals*. For whatever other possible names, he called it a *granule*. This granule stands at a higher level over its constituents and has a command over them and can manipulate them.[1] We consider the granule as a more general concept than the fuzzy set. Although the fuzzy set is a way to deal with one specific form of granule, namely, linguistic variables, not all granules are linguistic variables and hence not of all them are fuzzy. For example, signs or symbols studied in semiotics can be another type of granule, but they may have a precise definition or meaning and are not fuzzy.

In this article, we shall argue that granularity, in its broad sense, bears a close relation not just to the concept of aggregation in economics, but more to decision making and policy-making in economics. In a nutshell, granules are what make our decisions simple and efficient, having been coined the "fast and frugal heuristics" by Gerd Gigerenzer [21]. Their formation, development, and evolution are what enables human agents to effectively deal with the complex environment surrounding them.

With this in mind, we shall provide a comprehensive review of the literature which all points to the significance of granules as elementary units of information processing. These include Barry Schwartz's paradox of choice [44], George Miller's magic number 'seven' or chunking [35, 2, 45], and Herbert Simon's hierarchial modularity [4]. The use of granules in decision marking has far-reaching implications, as demonstrated in recent studies on the behavioral foundations of public policies, such as Gerd Gigerenzer's fast and frugal heuristics [21] and Richard Thaler and Cass Sunstein's 'nudges' [48]. We also show that our understanding of human-subject experiments can be dramatically different by using or not using the idea of

[1] A typical example is the fuzzy set with its membership function. Through the membership function, the constituents of the fuzzy set are *coordinated* in such a way that the set, as a whole, can be presented.

granularity in modeling their behavior [12]. In fact, granularity has been largely ignored in experimental studies involving adaptive artificial agents. Many algorithms for the designs of adaptive artificial agents, such as reinforcement learning or evolutionary computation, when applied to mimic or to replicate human-subject behavior, are often ignorant of the idea of granularity. We, however, demonstrate some exceptions in agent-based economic models which do take granularity into account when designing their artificial agents.

The rest of the paper is organized as follows. Section 2 reviews the theoretical foundations of granules. Section 3 shows their significance in regard to individuals' decision making or institutions' policy-making. Section 4 reviews the use of the idea of granularity in recent economic models of learning and adaptation either when applied to human subjects in the context of laboratory experiments or when applied to artificial agents in the context of agent-based computational economics. Section 5 gives the concluding remarks.

2 Social Science Theory of Granularity

2.1 Granularity of the Set of Alternatives

The three essential pillars of microeconomics are the utility function (objective function), the set of alternatives, and, finally, choice-making. With the publication of their magnum opus "*Theory of Games and Economic Behavior*" in 1944, John von Neumann and Oskar Morgenstern introduced to economists a paradigm to explicitly structure the three pillars, which is known as the expected utility maximization paradigm. This paradigm has repeatedly dominated the mainstream economics already for half a century. However, the reality of this paradigm has been constantly questioned in economics; accordingly, the structure of the three pillars has been incessantly given different conceptualizations. Among the three, maybe the least addressed one is the *set of alternatives* or, more specifically, the *cardinality* or the *granularity* of the set. While Barry Schwartz's influential book "*The Paradox of Choice*" [44] has already raised the possibility that a proper choice mechanism may not exist when the set of alternatives is too large or too fine, the granularity issue is still largely ignored in economic models of decision making.

2.1.1 The Choice Overload Hypothesis

The paradox of choice originates from a series of human-subject experiments which address the behavior related to choice conflicts, choice aversion or choice deferral. Obviously, in this situation, the subject is not well motivated to make a choice and, instead, prefers an indefinitely longer procrastination or simply not to make a choice. In the literature, the paradox of choice is formally known as the *choice overload hypothesis*. The hypothesis basically says that "an increase in the number of options to choose from may lead to adverse consequences such as a decrease in the motivation to choose or the satisfaction with the finally chosen option" ([43], p.73). The choice overload hypothesis was first proposed by Sheena Iyengar and Mark Lepper [31], and they also tested this hypothesis with a series of three experiments.

In their series of choice experiments, Iyengar and Lepper distinguished the designs with psychologically manageable numbers of choices (limited-choice condition), say, six, from the designs with psychologically excessive numbers of choices (extensive-choice condition), say, twenty-four. In their famous jam promotion experiment, different numbers of jam jars were displayed in two separate places (tables) in a supermarket, one with six different types of jam and one with twenty four different types of jam. They found that while the 24-jam table was able to attract more shoppers than the 6-jam one, it did not successfully beef up their purchasing willingness. In fact, only 3% of the shoppers at the 24-jam table subsequently purchased a jar of jam, whereas 30% of the shoppers at the 6-jam table did that.

In their two additional experiments, this 'more is less' result was also confirmed. In one case, students who were offered more topics (30) to write their essays for an extra credit did not show a higher interest to do so than students who were given fewer topics (6); and for those who actually did so did not perform better as opposed to that of the 6-choice group, in terms of the quality of the essay. In the other case, subjects were either given 6 or 30 different chocolates to choose. It was found that subjects with 30 choices might initially be more cheerful with this large assortment, but the choice process turned out to be difficult and frustrating and the result was that they were often not satisfied and felt regretful.

Since the freedom to choose is a cornerstone of any democratic society, the choice overload hypothesis does lead to an upheaval among academics and the public, which prompts more follow-up studies. Scheibehenne, Greifeneder, and Todd [43] provided a meta-analytic review of 29 articles (published and unpolished) with 50 experiments, from 2000 to 2009, involving 5,036 subjects. Among these 50 experiments, the minimum number for the limited-choice condition was 3, whereas the maximum number for the extensive-choice condition was 300. Using a random effecte model, they found that the results were mixed, neither supporting the choice overload hypothesis ("more is less") nor its opposite ("more is better").

In their meta-regression analysis, they, however, tended to suggest that the experimental results may be sensitive to some control variables pertinent to the design. Among them, maybe the most important one is the *experience effect*. If the subject is very experienced with the choice problem presented to them, for example, living in a town for years and having sufficient time to know all the restaurants around, he/she may not have a hard time choosing a place for lunch. In this situation, more choice can be better. In fact, this control variable was carefully fine-tuned when Iyengar and Lepper [31] initialized this line of research.

> In addition, to provide a clear test of the choice overload hypothesis, several additional methodological considerations seemed important. On the one hand, to minimize the likelihood of simple preference matching, *care was taken to select contexts in which most participants would not already have strong specific preferences*. (Ibid, p. 996; Italics added)

This may justify the use of exotic products in testing the choice overload hypothesis. What underlies the experience effect or the familiarity effect is the information regarding each alternative and the mechanism used to process the supplied information. The latter is further related to how the information is presented to the subject, i.e., the structure of the information, to which we now turn.

2.1.2 Characteristic Analysis

While making a choice, subjects may have to ask how this alternative is different from others. If their difference noticeably lies in one dimension, for example, 100 baskets containing different numbers of peanuts, then the choice overload issue may not happen because consumers can at least identify exactly what they want, for example, the basket with the maximum number of peanuts. Nonetheless, each alternative may have a number of attributes and they may differ in each attribute. This may cause the information required to distinguish them overwhelming and make a selection hard. Of course, issues can become simpler if these attributes are not presented in a wide flat, but can be endowed with a hierarchical structure.

In consumer theory, Kelvin Lancaster pioneers a different approach to the choice problem, called *characteristic analysis* [29]. In characteristic analysis, commodities (for example, different brands of toothpaste) are characterized by their attributes (characteristics) and the density (quantity or the quality) of those attributes. When presented with a set of alternatives, consumers search for the commodity which is closest to their desired attributes after taking into account the price they are required to pay. Lancaster further assumed that there is a hierarchical structure of these characteristics [30]. From an information processing viewpoint, this hierarchical structure enables decision makers to have a sequential decision process to deal with complex choice problems.

> A decision process in which a choice involving a *restricted number of parameters* is made, after which a further choice is made from *another restricted set of parameters*, and so on down the sequence, is necessarily hierarchical unless it is purely random. The ordering of the hierarchy determines which set of parameters is considered *first*, *second*, and on through the sequence. (Ibid, p.50; Italics added)

The implication of the quotation above is that the assortment structure is an important control variable while testing the choice overload hypothesis. Findings related to this observation are summarized well in [43]. For example, Mogilner, Rudnick, and Iyengar [36] found that an increase in the number of alternatives decreased satisfaction only if the alternatives were not displayed in categories.

To sum up, the paradox of choice, by and large, may exist only as a transition process as a short-term phenomenon.[2] Although humans are limited in their cognitive capacity, they can learn, adapt and hence cope with complex decision problems by developing decision heuristics to simplify hard choice problems. This adaptation process is, in fact, a granulation process. This granulation process can happen for both suppliers and consumers. For the suppliers, when a large number of options

[2] See [24] for some related discussions.

are displayed, these options will be ordered, categorized, grouped, and be given a hierarchical structure. For the consumers, they can take advantage of the given hierarchical structure or search with their own heuristics, for example, the elimination strategy, to sequentially reduce the search space and locate what they really want.

While the granulation process in a hierarchical manner is a way to escape from the paradox choice, how is the degree of granularity at each level determined? Why intuitively must it have a *coarse* division instead of a fine division at each level, as Lancaster has pointed out "*a restricted number of parameters*"? We shall address this question in the next section.

2.2 Cognitive Foundation of Granulation

2.2.1 Chunking and Magic Numbers

In this section, we try to examine the psychological foundation of granulation in the hierarchical form. A classic work which one cannot afford to miss is Miller's famous number *seven* [35]. Miller (1956) [35] is a celebrated contribution to psychology in the discussion of *short memory capacity* or *working memory capacity*. In this regard, it is about the number of items that an individual can discriminate or is about the capacity to remember information over very short periods of time, say, seconds. Based on a few experiments that he reviewed, Miller concluded that most people can correctly recall about 7 ± 2 items. This is the origin of the magic number seven.

For the purpose of this chapter, the significance of this work [35] is three-fold. First, it shows that through the granulation process a human can increase his memory span. In other words, granulation can be understood as a psychological process to enhance humans' capability to deal with a complex environment characterized by a large amount of information. This immediately brings us back to the early discussion of humans' capability to deal with the paradox of choice (Section 2.1). Second, it shows that the granulation process proceeds in a hierarchical manner. Third, while without being given an exact definition, the linguistic variables, as we shall see below, play a pivotal role in the granulation process. Hence, even though we argue in the very beginning of the chapter that a fuzzy set is only a special form of granulation, due to the heavy reliance on linguistic terms in the granulation process, the fuzzy set is clearly indispensable to the development of a general theory of granular computing.

Without rephrasing what he actually said and hence not distorting what he actually meant, we shall use two quotations directly from Miller's article to point out these connections. About the granulation process per se, he said the following.

> In order to speak more precisely, therefore, we must recognize the importance of grouping or organizing the input sequence into units or chunks.... In the jargon of communication theory, this process would be called *recoding*. The input is given in a code that contains *many chunks with few bits per chunk*. The operator recodes the input into another code that contains *fewer chunks with more bits per chunk*. There are many ways to do this recoding, but probably the simplest is to group the input events, apply *a new name* to the group, and *then remember the new name rather than the original input events*. ([35], p.93; Italics added.)

Chunking is probably the most influential idea we learned from Miller's studies [35]. The quotation above makes a distinction between *items* and *chunks*. With this distinction the granulation process can be regarded as a transition process from *many chunks with few bits per chunk* (items) to *fewer chunks with more bits per chunk* (chunks). This transition process is simply a process of *information compression*. While Miller's study was conducted in the middle of the 1950s, almost a decade earlier than the advent of *algorithmic information theory*, independently founded by Andrey Kolmogorov, Ray Solomonoff, and Gregory Chaitin in the mid-1960s, the idea of information compression as formations of chunks is already in the paper, as we quoted above. This helps clarify the subsequent discussions on *what exactly the magic number is*.

In their recent article, Fabien Mathy and Jacob Feldman [34] reconcile two versions of the magic number using a notion of *Kolmogorov complexity* and incompressibility. The two versions refer to the original seven (7 ± 2) and the later version of four (4 ± 1) [18]. Mathy and Feldman [34] assert that four is the true capacity of short-term memory in *maximally compressed units*, while Miller's magic number seven refers to the length of an uncompressed sequence. This number, seven or four, gives us a cognitive reason for granulation. To use our limited cognitive capacity more efficiently in order to increase our memory span, we tend to harness individuals as granules (compress items into chunks), and as a maximum we are able to have three to five chunks *at a level*.[3] These magic numbers correspond well to the choice overload hypothesis, which seems to indicate that if options are not arranged into categories (not compressed into groups), then, when the number of options increases beyond a threshold (the magic number), our motivation to make a choice or the satisfaction resulting from the option chosen will decrease.

2.2.2 Hierarchical Structure of Granules

As to the hierarchical form of granulation, Miller emphatically involved the ideas of what is currently known as *encapsulation*, which will also be discussed in Section 2.3.

> In my opinion the most customary kind of recoding that we do all the time is to translate into a *verbal code*. When there is a story or an argument or an idea that we want to remember, we usually try to rephrase it "in our own words." When we witness some event we want to remember, we make a verbal description of the event and then remember our *verbalization*. Upon recall we recreate by *secondary elaboration the details* that seem consistent with the particular verbal recoding we happen to have made. (Ibid, p. 95; Italics added)

Based on the quotation above, *the details in secondary elaboration* seem to indicate what are *inside* the chunks. While Miller did not make the hierarchical or recursive structure of chunks explicit, the subsequent interpretations of Miller's work do notice these branches. For example, Baddeley raises the following question [2].

[3] As we shall see below, there will be a hierarchical structure of these granules.

> The situation is further complicated by the possibility of setting up hierarchical structures of chunks. If seven chunks can be held, *can each one be divided into seven subchunks?* Presumably not, because that would suggest that one can hold 49 chunks. Perhaps the number seven, itself, comes from chunking. ...My own view is that it is unlikely that the limit is set purely by the number of chunks, independent of such factors as *the degree to which material within each chunk is integrated as a result, for example, of prior learning*. (Ibid, p. 355; Italics added)

Although Baddeley [2] did not give a clear answer with regard to the hierarchical structure of chunks, he did correctly point out that it would be hard to count the number of chunks independently of *prior learning*. In the following section, we shall introduce another notion related to granularity from the complex system perspective. There we shall see the relationship between learning and the evolution of the hierarchical granular system.

2.3 Hierarchical Modularity

Our discussion in the previous section indicates that what matters for the working memory capacity may not solely just be a number, 7 or 4. The question which we should really ask is: *what is inside the chunks?*. According to algorithmic information theory, chunks are compressed messages like a program with *minimum description length* [34]. How are these compressions actually made? Do they rely on some other existing programs or building blocks to facilitate the compression? If so, where do these building blocks come from? Has learning anything to do with them? This series of questions leads us to a highly influential concept in complex systems, namely, *modularity*. In this section, we shall first briefly review Herbert Simon's original work on *modularity* [46], and we shall then use LISP programming and genetic programming to demonstrate the learning process as a development of a hierarchical modular structure.

2.3.1 Modularity

This section is inspired by Herbert Simon's work on *near decomposability* or *modularity* [46]. Modularity refers to a structural relationship between a system as a whole and the constituent components which can function as independent entities. The interactions of the elements within the same constituent component are strong; however, the interactions of elements across different constituent components are weak, but not zero. The latter property is known as *near decomposability*. The chunk or granule, as a collection of items that have strong associations with one another but have much weaker associations with other chunks concurrently in use, is a typical near decomposable system. As Simon has argued, near decomposability is a key to harnessing a possibly unbounded complex system.

Simon [46] was probably one of the most influential pioneers inspiring many follow-up works in various scientific disciplines [4]. In addition to near decomposability, Simon viewed *hierarchy* as a general principle of complex structures. He advocated the use of a hierarchical measure – the number of successive levels of hierarchical structuring in a system – to define and measure complexity; furthermore, he argued that hierarchy emerges almost inevitably through a wide variety of evolutionary processes, for the simple reason that hierarchical structures are *stable*.

In addition to the depth of a hierarchy, Simon also noticed the span or the width of a hierarchy at each level. He defined the *span of a system* as the number of subsystems into which it was partitioned. Although he did exemplify some hierarchies with large or even indefinitely large spans, the so-called *flat hierarchies*, his attention was mainly drawn to the hierarchies of *moderate* span. While he did not give an exact range for a moderate span, it is our conjecture here that the magic number discussed in Section 2.2.1 can pinpoint a reasonable niche.

2.3.2 LISP

One example which is useful for us to think about the connection between the hierarchies of moderate span and granules or chunks is the *symbolic system*. In a symbolic system, the elementary units are alphabets (symbols). Using the grammar applied to the system, one can, syntactically correctly, generate words, sentences, books, and volumes. Each of these generated objects can be a granule or a chunk at different levels of a hierarchy of moderate span. Each of them, in an encapsulated form, may have some degree of independence, and can be reused as a chunk of other hierarchies.

In computer science, this is basically what *formal language theory* is about. Some computer languages clearly demonstrate this hierarchical structure, for example, LISP.[4] Each LISP program, regardless of its size, as a whole, is a *list*.[5] However, it may have other (sub)lists as its constituents, and each of them is also an

[4] LISP stands for List Processing, which is a high-level computer language invented by John McCarthy (1927-2011) in 1958 at MIT as a formalism for reasoning about the use of certain kinds of logical expressions, called recursion equations. This language is strongly motivated as a practical implementation of the λ calculus or the recursive function theory developed in the 1930s by Alonzo Church (1903-1995) and Alan Turing (1912-1954). See [1] for details.

[5] LISP S–expressions consist of either *atoms* or *lists*. Atoms are either members of a *terminal set*, that comprise the data (e.g., constants and variables) to be used in the computer programs, or they are members of a *function set* that consists of a number of prespecified functions or operators that are capable of processing any data value from the terminal set *and* any data value that results from the application of any function or operator in the function set. Lists are collections of atoms or lists, grouped within parentheses. In the LISP language, everything is expressed in terms of operators operating on some operands. The operator appears as the left-most element in the parentheses and is followed by its operands and a closing (right) parenthesis. For example, the S-expression $(+ X \; 3)$ consists of three atoms: from the left-most to right-most they are the function "+", the variable X and the constant 3. As another example, $(\times X \; (- Y \; 3))$ consists of two atoms and a list. The two atoms are the function "$\times$" and the variable "X," which is then followed by the list $(- Y \; 3)$.

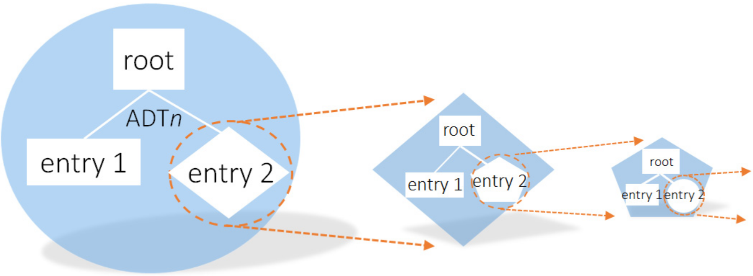

Fig. 1 Hierarchy of Chunks and Automatically Defined Terminals

This figure exemplifies an example of a hierarchy of chunks (granules or modules) through automaticaly defined terminals. Each chunk (granule or module), as represented by a LISP parse tree, has a span of two (two terminals, entry 1 and entry 2). However, each entry itself is also a chunk and has a span of two. This *self-similarity* can recursively go on and on for many levels.

independently implementable program. The macro list can been encapsulated and reused as a subroutine by other programs. This is what John Koza called the *automatically defined functions* (ADFs) [28], or alternatively, automatically defined terminals (ADT) [17] (see Figure 1 for a demonstration). Each automatically defined function or terminal can contribute to information compression as we mentioned in Section 2.2.1, i.e., they help reduce the otherwise more lengthy messages (codes) into smaller ones. Hence, they are chunks used to construct other chunks (see Figure 2 for an illustration).

By presenting all these chunks together, we can see that there is an underlying timeline as indexed by the generation number in Figure 2. To obtain more hierarchical chunks, some low-level chunks have to be developed first, such as ADT8, ADT9, ADT10, and ADT12 (Figure 2). In other words, there is a time order connecting these chunks, for example, ADT12 precedes ADT18 and ADT19. This timeline realized into the real world is, in effect, what from George Miller to Alan Baddeley has been called *learning*. This shows why Miller's magic number seven is not that absolute, because, at most, it only gives us the size of the span at each level. What really matters is the depth of the hierarchy, which is a consequence of learning.

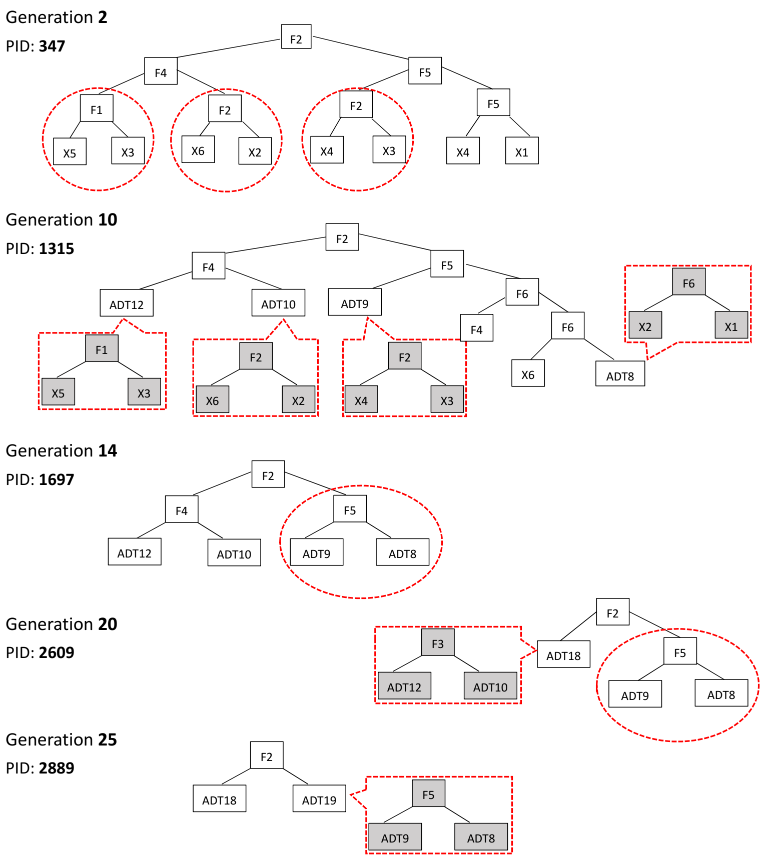

Fig. 2 Information Compression through Existing Chunks and Learning

This figure demonstrates a development process of a hierarchy of chunks (granules, modules). Here, some simpler automatically defined terminals (in terms of depth) are developed first, and they are encapsulated and are used as chunks (building blocks) to form more complex hierarchy. However, by using the existing chunks, the later derived hierarchies, no matter how complex they are, have a depth of no more than two.

3 Policy-Oriented Applications of Granularity

3.1 Granular Decision Making

The previous section, based on psychology, information theory, and complex science, has suggested a behavioral foundation of, as we shall coin in this section, *granular decision making*. The granular decision making refers to a hierarchical decision-making procedure which organizes an entire decision problem into several levels (layers). At each level, only a few granules are presented and need to be looked at; "coarser" granules are arranged at the higher level of the hierarchies, and "finer" granules are arranged at the lower level of the hierarchies. The decision maker starts at the first level (the topmost level) of the decision hierarchy, and stops at the level at which a decision has been made.

While we may have introduced a neologism, by no means do we want to claim its novelty. In fact, some familiar hierarchical decision frameworks have already existed, for example, the *analytical hierarchy process*.[6] However, this hierarchical decision framework is not necessarily applied in a way that is consistent with what is discussed in Section 2. For example, the number of granules required to be given at each level is not constrained by the 'magic number', and a decision cannot be made until all the information in the hierarchy has been processed.

This decision process may sound more systematic, comprehensive, and rigorous, but the fundamental question is: when can we trust the decision made through this process, and when should we better trust our gut feeling? Needless to say, the resolution of the issue is beyond the scope of the paper. However, we would like to point out that there are other hierarchical structures which are consistent with the essence of the granular decision making. Their existence and prevalence can been reviewed from two aspects. First, from the individual viewpoint (Section 3.2), we want to show how the granular decision making has been substantiated into practices to solve many of our daily life problems. Second, from the institutional viewpoint (Section 3.3), we want to show how public policy can become more effective or welfare enhancing if the *choice architecture* can be well taken care of in light of the granular decision-making framework.

3.2 Fast and Frugal Heuristics

The fast and frugal heuristics is pioneered and advocated by the German psychologist and behavioral economist, Gerd Gigerenzer. Over the years, he and the research that he had led at the Max Planck Institute have extensively studied how people actually make decisions when they are presented with complex, vague, and ill-defined problems. It is found that many heuristics, while they may look simple, can effectively solve problems in a fast and frugal manner, and even the results are sometimes better than those of deliberated complex models [23]. In the literature, there is a glossary of such heuristics available, such as the recognition heuristic, fluency heuristic, take-the-best heuristic, and one-good-reason heuristic.[7]

To give a highlight of these heuristics and their relationship with the granular decision making, one needs to know that Gigerenzer considered himself to be a behavioral economist following the legacy of Herbert Simon. Under the influence of Simon's notion of bounded rationality, behavioral economists characterize each

[6] Due to the space limitations, we do not intend to give a review of the large body on literature in the analytic hierarchy process. The interested reader is referred to [40, 25].

[7] For a review of these heuristics and many more, the interested reader is referred to [22].

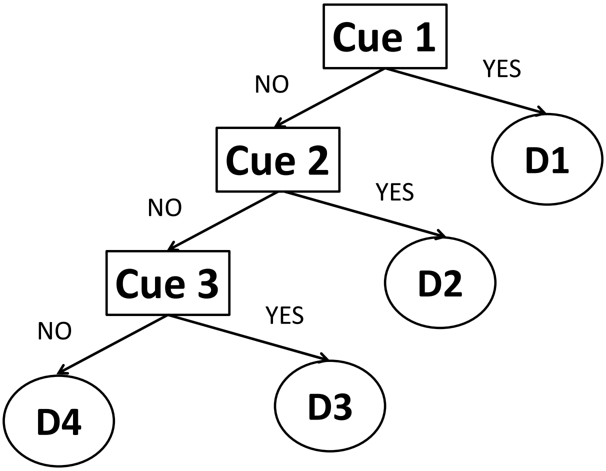

Fig. 3 Fast and Frugal Heuristics in the Form of Decision Trees

decision process with three main stays, namely, a *search rule*, a *stopping rule*, and a *decision rule*. A simple but abstract example is the *alphabet heuristic*.

To use the alphabet heuristic, each alternative is presumably characterized by a 'name' (a series of letters). By means of this heuristic, we search through the name of alternatives letter by letter in reading direction, and assess each letter's position in the alphabet (the search rule). If the letters in the first position differ among the alternatives, we only keep those whose letters appear the earliest in the alphabet, and remove the rest (the decision rule). If there is more than one left, we move to the second position, and then ditto the above procedure until there is only one left (the stopping rule). Then only the remaining one is chosen (the decision rule).

The alphabet heuristic has long been known as *lexicographic ordering* in economics. Its lexicographic structure is generally shared by the aforementioned heuristics, and is consistent with the granular decision making in the following regards. First, it has a hierarchical structure, first letter, second letter, and so on. Second, each letter actually has only two granules (a span of two), "the leading position in the alphabetic order" or "the positions behind it". Third, to make a decision, it is not necessary to go through all levels (or to the last level), as long as a decision can be made at an earlier stage. Fourth, learning plays a role in the formation of this hierarchy. In the case of the alphabet heuristic, one still needs to decide the alphabetical order: should it be 'ABC' or 'CBA'. Psychologists call these *cues*. In principle, one should prioritize those cues which are more informative or discriminating; for some decision problems this order of cues can be determined easily, but for some decision problems this order requires experience and learning.

The other hierarchical structure which is frequently used in machine learning, but less in psychology or behavioral economics, is *decision trees*. In fact, the fast-and-frugal heuristics reviewed above can be regarded as a special kind of decision tree, a binary decision tree as shown in Figure 3. In structure, it is much simpler than the general decision trees. It has only two nodes at each level, and at least one of the two is a terminal node. Because of this characteristic, each cue will appear only once in the tree. However, for a general decision tree, the number of nodes at each level varies, depending on the number of attribute values (not necessarily binary), and each cue can appear more than once as long as it does not repeat itself in the same subtree emanating from itself. Finally, the general decision trees are normally built using statistical algorithms, such as the entropy maximization algorithm, in the data mining context. Hence, learning in this situation is more formal and data-oriented; less depends on personal experiences or memory. Despite these technical differences, general decision trees, by and large, are also a kind of granular decision making.

3.3 Nudges

> Social science research reveals that as the choices become more numerous and/or vary on more dimensions, people are more likely to adopt simplifying strategies. The implications for choice architecture are related. As alternatives become more numerous and more complex, choice architects have more to think about and more work to do, and are much more likely to influence choices (for better or for worse). ([48], p. 95)

In Section 2.1.2, we have already mentioned that if options offered to decision makers can be well structured, then choice overload can be alleviated. An implication of this is then the way in which we present the options may affect what is chosen or not chosen. This lesson has already become known as the *framing effect* [49] for three decades. It is about two essentially identical choices provided to subjects, but one is presented in a *positive* frame, say, a bonus, and the other is presented in in a *negative frame*, say, a penalty. Simply by these different "phrases" people will be led to make a different choice. In a sense, the attention drawn to the choice-architecture effect can be regarded as a more extensive awareness of this framing effect.

Before we proceed further, it will be useful to mention one recent influential book in behavioral economics, entitled *Nudges*, authored by Richard Thaler and Cass Sunstein [48]. In a nutshell, what a *nudge* does is to make a *default* option available, and to use it to lessen the choice burden of people and to enable them to make a good choice. Some studies have shown that, by properly designing and incorporating a default option, we can have a choice architecture that can nudge people towards good decisions on spending, saving, health care, borrowing patterns, and organ donations. One of the most illustrative examples of the default option is to make enrollment in 401(k) plans *automatic* for new employees with a form for *opting out* so that if they do not wish to save they would have to register the desire to opt out of the plans. This *opt-out* system with a reasonable default saving plan is likely to result in greater retirement savings than an *opt-in* system [3].

What is the relevance of granularity to defaults? The answer is that, based on what we discuss in Section 2, *defaults are a kind of granule*. Depending on how a default is articulated, it can mean many different things with fine details. Using Miller's term (Section 2.2.1), it is a *name* or a verbal expression used for the information compression purpose. Once after a name is given we do not look further into the details being encapsulated until it is necessary to do so. Therefore, defaults share the psychological, cognitive, information processing, and complexity nature of granules (chunks, modules). The contribution of this chapter is to make their connection with granules clear.

The purpose of this chapter is not to provide an economic analysis of defaults. However, as Madrian [33] has pointed out, the economic studies on defaults is still very limited. Obviously, the significance of defaults in public policy should not be overrated. They are not panaceas. First, to make them work, they have to be properly designed. In our early example, we can have 'pop-in' or 'pop-out' as the default, but the latter (the automatic enrollment) is more effective than the former in increasing the participation rate of the retirement saving plan. Second, the function of a default or defaults should be evaluated in the hierarchy in which they are placed. Choosing a default may not mean the end of the story. For example, after automatic enrollment in the saving plan, the policy-makers can still offer few, but not many, options for employees to choose from.[8] In this situation, the default option can be applied again, which may refer to an endorsed fund. Hence, the design should be continued through the branch emanating from a default node. These additional layers certainly may take better care of the heterogeneous needs of employees than just the one-layer simple default.[9] Hence, from the granularity point of view, good policy-making does not rest upon a single default granule, but a granular hierarchy in which different defaults may be present at different levels.[10]

4 Granularity in Economic Modelling

Based on Duffy [19] and Chen [10], the two major learning or adaptive algorithms used in agent-based modeling are reinforcement learning and evolutionary computation (genetic algorithms and genetic programming). While both of these algorithms have been modified and extended into different forms, little attention and effort have been directed toward the granularity issue involved in the operation of the algorithms.

[8] The consideration is mainly to avoid the paradox of choice. Iyengar et al. [26] show that expanding the number of funds offered for the 401(k) contribution plan will lower the participation rate, which drops from 75 to 60 percent when the number of funds offered increases from only 2 to 59 options.

[9] Caroll et al. [8] provide a similar discussion. They actually suggest applying an active choice instead of defaults when there is substantial heterogeneity in consumer preferences.

[10] In practice, Amazon, NetFlix and many of the like, have shown themselves to be the masters of choice architecture as their entire websites are geared to enabling consumers to make desirable choices and continue transacting with them. As to the tools available for choice architects, the interested reader is referred to [27].

4.1 Reinforcement Learning

Let us first look at reinforcement learning. From [39], to [20], and further to [5], we have already extended the simple reinforcement learning algorithm to its generalized version by taking into account various cognitive or mental considerations, such as memory, consciousness, and reasoning. Nonetheless, the most fundamental unit is the choice or the action being reinforced, and hence the set of alternatives is the starting point for any version of reinforcement learning. From the granulation viewpoint, a fundamental question is: to make reinforcement learning applicable to understand a certain class of adaptive behavior, is there a granulation constraint that is required to be satisfied with, for example, the size of the set of alternatives? Although this question has been asked in a few studies, we believe that the granularity issue has never been addressed explicitly in any application of reinforcement learning known to us.

4.1.1 Beauty Contest Game

We shall illustrate this point by referring to some applications of reinforcement learning models to human-subject experiments. The specific one considered by us is the *beauty contest game*, which is also known as the *guessing game.* Players in the *beauty contest game* compete with each other to win a prize by selecting a number between $[0,100]$. The prize is given to the player whose guessed number is closest to the *target number*, which is calculated by averaging all guesses that are then post-multiplied by a factor p, say, $p = 2/3$. With this parameter, the game is called *the p-beauty contest game*. This game implicitly requires each player to form his/her expectations of other players' expectations. If other players are doing the same thing, the game then suffers from the familiar *infinite regress problem*. Under the homogeneous rational expectations hypothesis, a Nash equilibrium will be reached where everyone chooses an equilibrium of *zero*, which is the result of 50 (the middle point between 0 and 100) post-multiplied by p^{∞}. However, the resultant beauty contest experiments have demonstrated great deviations from this game-theoretic prediction [37, 38]. The application of reinforcement learning to the repeated beauty contest game was initiated by [5]. In their study, reinforcement learning was applied to a set of 101 alternatives.

In other words, each number that the subject can guess is an independent choice, and by the operation of reinforcement learning in the end the strategies demonstrating larger actual or simulated effects will be chosen more frequently. However, Camerer and Ho [5] found a lack of ability in reinforcement learning for modeling guessing dynamics in the beauty contest game, compared to other noncooperative games which are surveyed in their article. This result was also replicated in [12]. One of the remedies suggested in [5] is to consider the learning when players sophisticatedly realize that other players are learning as well.

Sophistication is a central concept in the beauty contest game for producing level-k reasoning and it has been put into practice in [6]. Chen and Du [12] also considered level-k reasoning during the learning process, yet the motivation and operation is different. In order to explicitly represent the granulation constraint imposed in learning dynamics, they suppose that the basic granules used for learning are only the six level-k rules instead of the 101 numbers. The learning model with these granules seems to be more descriptive than the one [5] which uses individual numbers directly. This result may lead to a fundamental question concerning the applicability of reinforcement learning to the situation when a large number of many possible choices are presented.

4.1.2 Coordination Game

The granularity issue can also be found in the application of the reinforcement learning model to the experimental behavior in the coordination game. The analytical framework was initiated by Van Huyck and his colleagues [51]. Let s_i^j be the action j of individual i and $s_i^j \in S = [0,1]$. The game is defined by the following payoff function $\pi(\cdot)$,

$$\pi(s_i, s_{-i}) = 0.5 - |s_i - \omega M(s)(1 - M(s)| \tag{1}$$

where s_{-i} denotes the actions of other players, $s = (s_i, s_{-i})$ is one action combination, $M(s)$ is the median of s and $\omega \in (1,4]$. There are two strict Nash equilibria, $(s,M) = (0,0)$ and $(1-1/\omega, 1-1/\omega)$.

Van Huyck and his colleagues conducted the human-subject experiments for this game [50]. The action space is discretized into a finite set containing 101 actions such that $S = \{0, 0.01, 0.02, ..., 0.99, 1\}$. Subjects in the group of five enter each session playing repeatedly for a total of T periods, where $T = 40, 70$ or 75. This experimental data offers an opportunity for researchers, not only Van Huyck and his colleagues, to study the learning behavior by fitting various kinds of learning models [16, 42, 50]. Reinforcement learning serves as a common algorithm fitted in all of those studies.

Some authors have well recognized that the size of the set of alternatives might be too large to be true, and they further suggested that the *similarity* among strategies should be used to simplify the decision problem [16, 42]. They assume that the strategies are ordered numbers and $s^j < s^k$ if $j < k$. The common similarity function being considered in those studies is the *Bartlett similarity function*, which supposes that the degree of similarity decreases linearly as the distance between the chosen strategy s^k and other strategies s^j, $\forall j \neq k$, increases. The similarity function is defined as follows:

$$f(s^j, s^k; h) = \begin{cases} 1 - |j-k|/h & \text{if } |j-k| \leq h, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

where h determines the $h-1$ unchosen strategies on either side of the chosen strategy. The updating rule for the propensity or attraction of strategy j, $R^j(t)$, is governed by

$$R^j(t) = rR^j(t-1) + f(s^j, s^k; h)\pi_k(t) \tag{3}$$

where $\pi_k(t)$ is the payoff of chosen strategy k in round t.

The incorporation of strategy similarity into the operation of reinforcement learning grasps the idea that the neighborhood of actions or strategies (within a granule) should be dealt with *as a whole rather than individually*. On the one hand, we might argue that what makes a granule is not explicitly defined here. On the other hand, if we could consider each chosen *case* as a centroid of a granule, the number of granules might grow throughout the action space as the game is played repeatedly. In the extreme case, the number of granules will be equal to the number of strategies bringing in another granularity issue needing to be solved.

4.1.3 Market

Previous mentioned applications of reinforcement learning, including the guessing game and the coordination game, are both examples of strategic games which are simpler. Market participants usually encounter a more complex learning task. Even though the rules and settings can be controlled and simplified in the laboratory, the time spent for instruction is usually longer and the practice session is always required to insure their comprehension of the *experimental markets*. It is also more difficult to identify what or to characterize how the subjects learned. Give those complexity , it seems that the issue of granularity, even though not been mentioned explicitly, was naturally and inevitably considered when applying reinforcement learning to model the adaptive process in experimental markets.

Chen and Hsieh [13] apply reinforcement learning to an order book-driven experimental prediction market. They assume that the mental representation of the decision problem has been simplified to *the choice of the intensity of limit order submission*. It is found that the intensity of limit order submission is positively correlated with the profit earned. They further assume that a ternary choice problem to reduce the possible dimensions of learning. In particular, the reinforcement mechanism takes effect on updating the propensity of three alternatives including increase the use of limit orders, decrease it, or keep it unchanged. In fact, it assumes what really matter is not a precise degree of intensity but a rough class with a coarse granule.

The second example is an application of generalized reinforcement learning to the adaptive process in *bilateral call market* [7]. In this market, a buyer and a seller are randomly matched each other and privately informed about their values or costs of the goods. They submit bid or ask anonymously and a trade is made at the midpoint when the buyer's bid is higher than the seller's ask. Subjects need to learn how to submit a desirable bid/ask conditional on his private value/cost on hand. For a buyer i, the strategy s_i being considered is the combination of the assigned private value v_i, $v_i \in [0, 200]$, and his submitted bid b_i, which will be determined by the choice of markdown ratio from 0 to 1. In order to implement reinforcement learning model, Camerer, Hsia, and Ho [7] discretize v_i into 10 equal intervals and 16 evenly spaced markdown ratios from 0 to 75%. We might consider this discretization as necessary although the resulting strategy space with 160 alternatives seems to be far from the range of the 'magic number' (Section 2.2.1).

In addition to discretization, Camerer and his colleagues introduced a parameter τ into the generalized reinforcement learning model to characterize the *similarity-based generalization of learning*, which shares the common motivation with other work introduced in Section 4.1.2 [16, 42]. The τs are different for each strategy and are sensitive to the distance between the strategy being considered and the chosen strategy. The definition of τ is given as follows:

$$\tau = e^{-\psi|v-v_i(t)|-\omega|b-b_i(t)|} \tag{4}$$

where $v_i(t)$ is the realized private value and $b_i(t)$ is the chosen bid of buyer i at time t. For the chosen strategy $s_i(t) = (v_i(t), b_i(t))$, τ is equal to its maximum 1 because $v - v_i(t)$ and $b - b_i(t)$ are equal to zero. On the other hand, for the non-chosen ones, τ can be close to zero depending on the magnitude of ψ and ω. The influence of parameter τ is so overwhelming that it can turn on/off the updating process of strategy attraction[11]:

$$A_i(t) = \frac{\phi^{\tau} N(t-1) A_i(t-1) + \tau \pi_i(b_i, v_i, v_{-i}(t))}{\phi^{\tau}(1-\kappa)^{\tau} N(t-1) + \tau} \tag{5}$$

Notice that when $\tau = 0$, $A_i(t)$ is equal to $A_i(t-1)$ indicating all adaptive mechanisms/parameters take no effect.

We have discussed and commented the incorporation of strategy similarity into the operation of reinforcement learning in Section 4.1.2. We would like to make a final remark on it by first quoting a paragraph from [7].

> Similarity-based learning is also arguably a "cognitively economical" heuristic because scarce attention is allocated where it is likely to be most useful - namely, in the vicinity of the current valuation and bid. (Ibid, p. 257)

In other words, subjects may focus only on the chosen strategy due to limited attention, and the other unchosen strategies are reinforced by the "spillover effect" generated from an automatic, unconscious process of having less requirement of cognitive load. This attention-based similarity learning further alleviate the dimensionality problem *when updating the attraction levels*. However, the *further calculation of choice probability* is still based on the strategy attractions of 160 finer granules (after disctretization), which implicitly assumes that subjects, in each round of their choice making, are still required to distinguish the attraction of each of this overwhelming number of alternations. Hence, we arrive the same conclusion that we have indicated in Section 4.1.2: this similarity device alone is not sufficient enough to solve the granularity issue.

---

[11] In order not to distract the attention of the readers, more details of the updating rule refer to [7].

4.2 Evolutionary Computation

Evolutionary computation involves *populations* of programs, strings of symbols, or candidate solutions. In agent-based economic models, sometimes the entire population is used to represent a whole society in a one-to-one mapping, known as *population evolutionary computation*; sometimes, the whole population is used to represent a single individual, known as *individual evolutionary computation*. Economists have frequently used social learning and individual learning to distinguish the former from the latter in their agent-based modeling [52]. The granularity issue does not appear in the former case where each individual agent is only associated with one single strategy and hence no choice. However, it may occur in the case of individual learning where each agent is endowed with a population of strategies and each time upon making a decision or taking an action he/she needs to choose one from them. Will they encounter the issue of the paradox of choice when the population size gets big? Will they need a proper choice architecture to deal with the large population of options? Will they need a search strategy to limit their search efforts. This issue, to the best of our knowledge, has never been raised.

In early days when evolutionary computation had only just been introduced to economics and applied to build agent-based economic models, the idea of ‘more is less’ had not even been formally proposed. Barry Schwartz’s book was available only after 2003, and before 2003 economists were not fully aware of the issue of choice overload. On the other hand, one may argue that the choice overload problem does not apply because in most applications alternatives differ in a single dimension only, namely, fitness or performance. In this case, like the early example which we mention, one only needs to search for the basket with the maximum number of peanuts. Without too many additional attributes, making comparisons among these alternatives is straightforward. Hence, even though some economists may still be concerned with the parameter of population size, they are motivated by different reason [32].

After the middle of the 2000s, some economists started to notice the connection between cognitive psychology and population size, and even Miller’s magic number has also been cited as a reference to determine the population size [9]. A research agenda developed along this research line is to use agent-based modeling to address the effect of cognitive capacity on earnings performance under a competitive market environment, and population size is used as a proxy variable for cognitive capacity [14, 15]. Even up to this step, whether or not our artificial agents are overloaded with too many choices is not a concern.

Exceptions do, however, exist. In a follow-up study of the Santa Fe artificial stock market, Tay and Linn [47] state:

> [S]ome might question whether it is reasonable to assume that *traders are capable of handling a large number of rules* for the mapping of market states into expectations, each with numerous conditions, ... We show this by allowing agents the ability to *compress information into a few fuzzy notions* which they can in turn process and analyze with fuzzy logic. (Ibid, p. 322; Italics added.)

Indeed, the above quotation echoes well with the discussion in Section 2.2.1 as well as with the quotation which we cite from Lofti Zadeh at the beginning of the chapter. It makes little sense of the large set of alternatives when they are presented in continuous fashion by rational numbers. It would be more sensible to add a *name* to some of them and treat them *as a whole* (as a granule). In this way, the effective strategy space can be substantially reduced. What Tay and Linn did was to use linguistic variables to group the strategies in such a way that one will not overburden their decisions with just number-crunching trivialities.

A similar approach has also been adopted by Chen and Chie [11] in an agent-based lottery market. In this study, they want to model agents' decision rules regarding lottery participation, and the decision rule depends on the lottery market condition characterized by the Jackpot size. Like Tay and Linn [47], they did not directly use rational numbers to define the market condition; instead, the fuzzy sets and linguistic variables were applied to give different states of the market. The Sugeno style of fuzzy rules was then applied to form the lotto participation rules.

5 Concluding Remarks

Generally speaking, the concept of granularity has been largely ignored by economists in their models of decision marking. However, thanks to the recent interdisciplinary research joining economists and psychologists, which enables us to realize the possible implications of Shannon information capacity for the human mind and to cast doubt on the fundamental assumption of economics: more options can only do good, or at least no harm. This gives us a psychological foundation of the granule (from fuzzy mathematics), chunk (from cognitive psychology), or module (from complex science) as an elementary unit of information processing and decision making. Moreover, through learning and constant information compression, these elementary units are accumulated and arrayed in a hierarchical form so that our memory capacity can be less constrained. Experimental studies with human subjects also find that subjects' learning behavior can be well captured by the reinforcement learning model if the size of the set of alternatives is restricted to a small number, say, 6, and not 101 [12]. In other words, when agents are presented with a large number of individual options, they group them as granules and mentally work with these granules (as a whole) rather than with the constituent individuals.

This chapter just serves as the beginning of a new research direction for both experimental economics and agent-based computational economics. As for the former, it prompts us to design laboratory experiments or field studies to address the effect of the number of options on choice behavior, with or without the choice architecture. The role of defaults and their possible forms can be closely scrutinized in these experiments. As for the latter, it motivates us to apply the granular information processing scheme to modeling artificial agents. In this case, the default at the very top of the hierarchy may mean the *status quo heuristic*.

The *status quo heuristic* has been well studied in psychology and behavioral economics [41]. It means that in many choice situations, people value the *status quo*, and will forego the opportunity to switch to an alternative unless it is really necessary to do so. The status quo heuristic has been applied in agent-based models in different forms, the threshold width in the threshold model, the intensity of choice in the stochastic choice model, and even the steady-state replacement used in evolutionary computation. To some extent, all try to give a weight for the inertial tendency. However, few have explicitly acknowledged the underpinning psychological force.

Finally, given the hierarchical structure of granules (or granular information processing), it is important to take a further look at the neologism. If granules have to be named verbally, then the evolution of 'names' (still using Miller's term) should give us a footprint of a hierarchy from its primitives to the holistic one. For example, in considering 'happiness' as a granule, what inside this is granule may differ over time; at one time, it can mean homeland security, and at another time, it may mean finding a real job. Therefore, it opens a new navigation in the literature.

Acknowledgements. The authors are grateful to Ying-Fang Kao, Bin-Tzong Chie, and two anonymous referees for their suggestions made to the paper. All remaining errors are the authors' sole responsibility. The Ministry of Science and Technology grant MOST 103-2410-H-004-009-MY3 is gratefully acknowledged.

References

1. Abelson, H., Sussman, G.: Structure and Interpretation of Computer Programs, 2nd edn. MIT Press (1996)
2. Baddeley, A.: The magical number seven: Still magic after all these years? Psychological Review 101(2), 353–356 (1994)
3. Beshears, J., Choi, J., Laibson, D., Madrian, B.: The importance of default options for retirement saving outcomes: Evidence from the United States. In: Kay, S., Sinha, T. (eds.) Lessons from Pension Reform in the Americas, pp. 59–87. Oxford University Press, New York (2008)
4. Callebaut, W., Rasskin-Gutman, D. (eds.): Modularity: Understanding the Development and Evolution of Natural Complex Systems. MIT Press, Cambridge (2005)
5. Camerer, C., Ho, T.-H.: Experienced-weighted attraction learning in normal form games. Econometrica 67(4), 827–874 (1999)
6. Camerer, C., Ho, T.-H., Chong, J.-K.: Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. Journal of Economic Theory 104(1), 137–188 (2002)
7. Camerer, C., Hsia, D., Ho, T.-H.: EWA learning in bilateral call markets. In: Rapoport, A., Zwick, R. (eds.) Experimental Business Research, pp. 255–284 (2002)
8. Carroll, G., Choi, J., Laibson, D., Madrian, B., Metrick, A.: Optimal defaults and active decisions: Theory and evidence from 401(k) saving. Quarterly Journal of Economics 124, 163–1674 (2009)
9. Casari, M.: Can genetic algorithms explain experimental anomalies? An application to common property resources. Computational Economics 24, 257–275 (2004)

10. Chen, S.-H.: Reasoning-based artificial agents in agent-based computational economics. In: Nakamatsu, K., Jain, L. (eds.) Handbook on Reasoning-Based Intelligent Systems, pp. 575–602. World Scientific (2013)
11. Chen, S.-H., Chie, B.-T.: Lottery markets design, micro-structure, and macro-behavior: An ACE approach. Journal of Economic Behavior and Organization 67(2), 463–480 (2008)
12. Chen, S.-H., Du, Y.-R.: Heterogeneity in experience-weighted attraction learning and its relation to cognitive ability. Working paper, AI-ECON Research Center, National Chengchi University (2014)
13. Chen, S.-H., Heieh, Y.-L.: Reinforcement Learning in Experimental Asset Markets. Eastern Economic Journal 37, 109–133 (2011)
14. Chen, S.-H., Tai, C.-C.: The agent-based double auction markets: 15 years on. In: Takadama, K., Cioffi-Revilla, C., Deffuant, G. (eds.) Simulating Interacting Agents and Social Phenomena, pp. 119–136. Springer (2010)
15. Chen, S.-H., Yu, T.: Agents learned, but do we? Knowledge discovery using the agent-based double auction markets. Frontiers of Electrical and Electronic Engineering (FEE) in China 6(1), 159–170 (2010)
16. Chen, Y., Khoroshilov, Y.: Learning under limited information. Games and Economic Behavior 44, 1–25 (2003)
17. Chie, B.-T., Chen, S.-H.: Non-price competition in a modular economy: An agent-based computational model. Economia Politica: Journal of Analytical and Institutional Economics XXX, 149–175 (2013)
18. Cowan, N.: The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences 24, 87–185 (2001)
19. Duffy, J.: Agent-based models and human subject experiments. In: Tesfatsion, L., Judd, K. (eds.) Handbook of Computational Economics: Agent-based Computational Economics, vol. 2, pp. 949–1011. Elsevier, Oxford (2006)
20. Erev, I., Roth, A.: Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy euilibria. American Economic Review 88, 848–881 (1998)
21. Gigerenzer, G.: Gut Feelings: The Intelligence of the Unconsciousness. Penguin Books (2007)
22. Gigerenzer, G., Gaissmaier, W.: Heuristic decision making. Annual Review of Psychology 62, 451–482 (2011)
23. Gigerenzer, G., Todd, P.: Fast and Frugal Heuristics: The Adaptive Toolbox. In: Gigerenzer, G., Todd, P. (eds.) Simple Heuristics That Make Us Smart, pp. 3–34. Oxford University Press, New York (1999)
24. Hutchinson, J.: Is more choice always desirable? Evidence and arguments from leks, food selection, and environmental enrichment. Biological Reviews 80(1), 73–92 (2005)
25. Ishizaka, A., Labib, A.: Review of the main developments in the analytic hierarchy process. Expert Systems with Applications 38, 14336–14345 (2011)
26. Iyengar, S., Huberman, G., Jiang, W.: How much choice is too much? Contributions to 401(k) Retirement Plans. In: Mitchell, O., Utkus, S. (eds.) Pension Design and Structure: Lessons from Behavioral Finance, pp. 83–95. Oxford University Press (2004)
27. Johnson, E., Shu, S., Dellaert, B., Fox, C., Goldstein, D., Haubl, G., Larrick, R., Payne, J., Peters, E.: Beyond nudges: Tools of a choice architecture. Marketing Letters 23, 487–504 (2012)

28. Koza, J.R.: Genetic Programming II–Automatic Discovery of Reusable Programs. The MIT Press, Cambridge (1994)
29. Lancaster, K.: A new approach to consumer theory. Journal of Political Economy 74(2), 132–157 (1966)
30. Lancaster, K.: Hierarchies in goods-characteristics analysis. Advances in Consumer Research 3, 348–352 (1976)
31. Lyengar, S., Lepper, M.: When choice is demotivating: Can one desire too much of a good thing? Journal of Personality and Social Psychology 79(6), 995–1006 (2000)
32. Lux, T., Schornstein, S.: Genetic learning as an explanation of stylized facts of foreign exchange markets. Journal of Mathematical Economics 41(1-2), 169–196 (2005)
33. Madrian, B.: Applying insights from behavioral economics to policy design. Annual Review of Economics 6, 663–688 (2014)
34. Mathy, F., Feldman, J.: What's magic about magic numbers? Chunking and data compression in short-term memory. Cognition 122, 346–362 (2012)
35. Miller, G.: The magical number seven plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, 81–97 (1956)
36. Mogilner, C., Rudnick, T., Iyengar, S.: The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. Journal of Consumer Research 35(2), 202–215 (2008)
37. Nagel, R.: A survey on beauty contest experiments: Bounded rationality and learning. In: Budescu, D., Erev, I., Zwick, R. (eds.) Games and Human Behavior, Essays in Honor of Amnon Rapoport. Lawrence Erlbaum Associates, Inc., New Jersey (1998)
38. Nagel, R.: Experimental beauty contest games: Levels of reasoning and convergence to equilibrium. In: Plott, C.R., Smith, V. (eds.) Handbook of Experimental Economics Results, vol. 1, ch. 45, pp. 391–410. Elsevier (2008)
39. Roth, A., Erev, I.: Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior, Special Issue: Nobel Symposium 8, 164–212 (1995)
40. Saaty, T.: How to make a decision: The analytic hierarchy process. European Journal of Operational Research 48, 9–26 (1990)
41. Samuelson, W., Zeckhauser, R.: Status quo bias in decision making. Journal of Risk and Uncertainty 1, 7–59 (1988)
42. Sarin, R., Vahid, F.: Strategy similarity and coordination. The Economic Journal 114, 506–527 (2004)
43. Scheibehenne, B., Greifeneder, R., Todd, P.: Can there ever be too many options? A meta-analytic review of choice overload. Journal of Consumer Research 37(3), 409–425 (2010)
44. Schwartz, B.: The Paradox of Choice: Why More Is Less. Harper Perennial (2003)
45. Shiffrin, R., Nosofsky, R.: Seven plus or minus two: A commentary on capacity limitations. Psychological Review 101(2), 357–361 (1994)
46. Simon, H.: The architecture of complexity. Proceedings of the American Philosophical Society 106(6), 467–482 (1962)
47. Tay, N., Linn, S.: Fuzzy inductive reasoning, expectation formation and the behavior of security prices. Journal of Economic Dynamics & Control 25, 321–361 (2001)
48. Thaler, R., Sunstein, C.: Nudge: Improving Decisions about Health, Wealth, and Happiness. Penguin Books (2008)

49. Tversky, A., Kahneman, D.: The framing of decisions and the psychology of choice. In: Wright, G. (ed.) Behavioral Decision Making, pp. 24–41. Plenum (1981)
50. Van Huyck, J., Battalio, R., Rankin, F.: Selection dynamics and adaptive behavior without much information. Economic Theory 33, 53–65 (2007)
51. Van Huyck, J., Cook, J., Battalio, R.: Selection dynamics, asymptotic stability, and adaptive behavior. Journal of Political Economy 102, 975–1005 (1994)
52. Vriend, N.: On two types of GA-Learning. In: Chen, S.-H. (ed.) Evolutionary Computation in Economics and Finance, pp. 233–243. Physica-Verlag, Heidelberg (2001)
53. Zadeh, L.: Fuzzy sets and information granularity. In: Gupta, M., Ragade, R.K., Yager, R.R. (eds.) Advances in Fuzzy Set Theory and Applications, pp. 3–18. North-Holland Publishing Company (1979)