



The Chaos Within Sudoku

Mária Ercsey-Ravasz¹ & Zoltán Toroczkai^{2,3}

SUBJECT AREAS:

STATISTICAL PHYSICS,
THERMODYNAMICS AND
NONLINEAR DYNAMICS

INFORMATION THEORY AND
COMPUTATION

MATHEMATICS AND
COMPUTING

PHYSICS

Received
7 August 2012

Accepted
26 September 2012

Published
11 October 2012

Correspondence and
requests for materials
should be addressed to
M.E.R. (ercsey.
ravasz@phys.ubbcluj.
ro) or Z.T. (toro@nd.
edu)

¹Faculty of Physics, Babeş-Bolyai University, Str. Kogalniceanu Nr. 1, RO-400084 Cluj-Napoca, Romania, ²Interdisciplinary Center for Network Science and Applications (iCeNSA), ³Departments of Physics, Computer Science and Engineering, University of Notre Dame, Notre Dame, IN, 46556 USA.

The mathematical structure of Sudoku puzzles is akin to hard constraint satisfaction problems lying at the basis of many applications, including protein folding and the ground-state problem of glassy spin systems. Via an exact mapping of Sudoku into a deterministic, continuous-time dynamical system, here we show that the difficulty of Sudoku translates into transient chaotic behavior exhibited by this system. We also show that the escape rate κ , an invariant of transient chaos, provides a scalar measure of the puzzle's hardness that correlates well with human difficulty ratings. Accordingly, $\eta = -\log_{10} \kappa$ can be used to define a "Richter"-type scale for puzzle hardness, with easy puzzles having $0 < \eta \leq 1$, medium ones $1 < \eta \leq 2$, hard with $2 < \eta \leq 3$ and ultra-hard with $\eta > 3$. To our best knowledge, there are no known puzzles with $\eta > 4$.

In Sudoku, considered as one of the world's most popular puzzles¹, we have to fill in the cells of a 9×9 grid with integers 1 to 9 such that in all rows, all columns and in nine 3×3 blocks every digit appears exactly once, while respecting a set of previously given digits in some of the cells (the so-called clues). Sudoku is an exact cover type constraint satisfaction problem² and it is one of Karp's 21 NP-complete problems³, when generalized to $N \times N$ grids⁴. NP-complete problems are "intractable" (unless $P=NP$)^{2,5} in the sense that all known algorithms that compute solutions to them do so in exponential worst-case time (in the number of variables N); in spite of the fact that if given a candidate solution, it takes only polynomial time to check its correctness.

The intractability of NP-complete problems has important consequences, ranging from public-key cryptography to statistical mechanics. In the latter case, for the ground-state problem of Ising spin glasses (± 1 spins), one needs to find the lowest energy configuration among all the 2^N possible spin configurations, where N is the number of spins. Additionally, to describe the statistical behavior of such Ising spin models, one has to compute the partition function, which is a sum over all the 2^N configurations. Barahona⁶, then Istrail⁷ have shown that for non-planar crystalline lattices, the ground-state problem and computing the partition function are NP-complete⁷. Since there is little hope in providing polynomial time algorithms for NP-complete problems, the focus shifted towards understanding the nature of the complexity forbidding fast solutions to these problems. There has been considerable work in this direction, especially for the boolean satisfiability problem SAT (or k -SAT), which is NP-complete for $k \geq 3$. Completeness means that *all problems* in NP (hence Sudoku as well), can be translated in polynomial time and formulated as a k -SAT problem, as shown for the first time by Cook and Levin². Namely, any problem in NP can be solved via a small number of calls to a k -SAT solver and a polynomial number of steps (in the size of the input) outside the subroutine invoking the k -SAT solver.

In k -SAT we are given N boolean variables to which we need to assign 1s or 0s (TRUE or FALSE) such that a given set of clauses in conjunctive normal form, each containing k or fewer literals (literal: a boolean variable or its negation) are all satisfied, i.e., evaluate to TRUE. Just as for the spin glass model, here we also have exponentially many (2^N) configurations or assignments to search.

In the following we treat algorithms as *dynamical systems*. An algorithm is a finite set of instructions acting in some state space, applied iteratively from an initial state until an end state is reached. For example, the simplest algorithm for the Ising model ground state problem, or the 3-SAT problem would be exhaustively testing potentially all the 2^N configurations, which quickly becomes forbidding with increasing N . To improve performance, algorithms have become more sophisticated by exploiting the *structure of the problem* (of the state space). Accordingly, now 3-SAT can be solved by a deterministic algorithm with an upper bound of $O(1.473^N)$ steps⁸. Here we will only deal with deterministic algorithms that is, once an initial state is given, the "trajectory" of the dynamical system is uniquely determined. Thus, we expect that the dynamics of those algorithms that exploit the structure of hard problems will reflect the complexity inherent in the problem itself. Complex behavior by *deterministic* dynamical systems is coined chaos in the literature^{9–13}, and thus the behavior of algorithms for hard problems is expected to appear highly irregular or chaotic¹⁴.

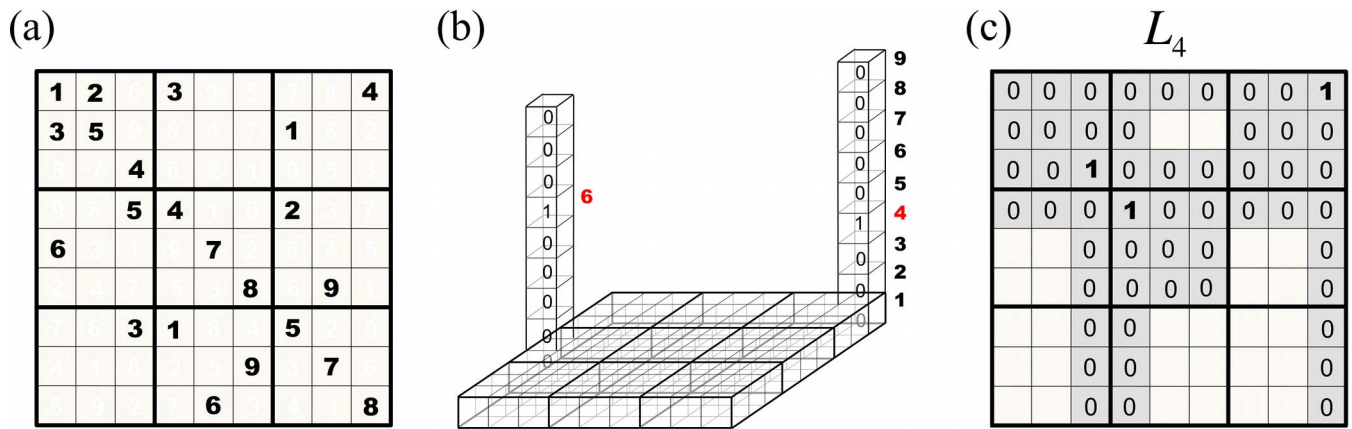


Figure 1 | Sudoku and its boolean representation. (a) a typical puzzle with bold digits as clues (givens). (b) Setup of the boolean representation in a $9 \times 9 \times 9$ grid. (c) Layer L_4 of the puzzle (the one containing the digit 4) with 1-s in the location of the clues and the regions blocked out for digit 4 by the presence of the clues (shaded area).

Although the theory of nonlinear dynamical systems and chaos is well-established, it has not yet been exploited in the context of optimization algorithms. One of the difficulties lies with the fact that most optimization algorithms are discrete and not easily cast in forms amenable to chaos theory methods. Recently, however, we have provided¹⁵ a *deterministic* continuous-time solver for the boolean satisfiability problem k -SAT using coupled ordinary differential equations (ODE) with a one-to-one correspondence between the k -SAT solution clusters and the attractors of the corresponding system of ODEs. This continuous-time dynamical system (CTDS) is in a form naturally suited for chaos theory methods, and thus it allows us to study the relationship between optimization hardness and chaotic behavior. Here we will focus only on solvable (satisfiable) instances, and thus the observed chaotic behavior will necessarily be transient^{11–13,16,17}. We need to emphasize, however, that the dynamical properties characterize both the problem and the algorithm itself. For this reason, one compares the dynamical properties across problems of varying hardness using the same algorithm. Nevertheless, since there are problem instances that are hard for all known algorithms, the appearance of transient chaos with long lifetime should be a universal feature of hard problems. It is also important to observe that transient chaos is not an $N \rightarrow \infty$ asymptotic behavior (where N is the size of the problem in terms of input or number of variables), but it appears for finite N , and thus measures of chaos can be used to characterize and categorize the hardness of *individual instances* of finite problems. To illustrate this, here we first map the popular 9×9 (hence finite) version of Sudoku into k -SAT, then we solve it using our deterministic continuous-time solver¹⁵. By analyzing the behavior of the corresponding trajectories of the CTDS we show the appearance of transient chaos when increasing the hardness of the Sudoku problems, and show that the level of hardness (taken from human ratings of the puzzles) correlates well with a chaotic invariant, namely the *lifetime of chaos* estimated as κ^{-1} , where κ is called the escape rate¹¹. We conclude with a discussion on algorithmic performance, dynamical properties and problem complexity.

Results

Sudoku as k -SAT. Because our continuous-time dynamical system¹⁵ was designed to solve k -SAT formulae in conjunctive normal form (CNF), we first briefly describe how Sudoku can be interpreted as a +1-in-9-SAT formula, and then how it is transformed into the standard CNF form. Further details are shown in the Methods section.

In a Sudoku puzzle we are given a square grid with $9 \times 9 = 81$ cells, each to be filled with one of nine symbols (digits) $D_{ij} \in \{1, \dots, 9\}$, i, j

$= 1, \dots, 9$ (with the upper-left corner of the puzzle corresponding to $i = 1, j = 1$). When the puzzle is completed, each of the columns, rows and 3×3 sub-grids (blocks partitioned by bold lines, Fig. 1a) must contain all the 9 symbols. Equivalently, all 9 symbols must appear once and only once in each row, column and 3×3 sub-grid.

To formulate Sudoku as a constraint satisfaction problem (CSP) using boolean variables, we associate to each symbol (digit) an ordered set of 9 boolean variables (TRUE = “1”, FALSE = “0”). The digit D_{ij} in cell (i, j) will be represented as the ordered set $(x_{ij}^1, \dots, x_{ij}^9)$ with $x_{ij}^a \in \{0, 1\}$, $a = 1, \dots, 9$, such that always one and only one of them is 1 (TRUE). In particular, $D_{ij} = a$ is set to be equivalent to $x_{ij}^a = \delta_{a,b}$, where $\delta_{a,b}$ is the Kronecker delta function. This way we have in total $9 \times 9 \times 9 = 729$ boolean variables x_{ij}^a , which we can picture as being placed on a 3D grid (Fig. 1b), with a corresponding to the grid index along the vertical direction, and hence a is the digit that is filling the corresponding (i, j) cell in the original puzzle. The corresponding horizontal 9×9 2D layer at height a will be denoted by L_a . Introducing the notion of such (horizontal) layers makes it easier to express the constraints of the Sudoku rules on its representation by 0-s and 1-s as described below. For example in the puzzle shown in Fig. 1a, we have $D_{1,9} = 4$. In the given vertical column the variable in the a^{th} cell (that is in layer L_a) is $x_{1,9}^a = \delta_{a,4}$ (shown as the boolean variable 1 filling the cell next to the digit 4 shown in red, in Fig 1b). This setup allows us to encode the Sudoku constraints in a simple manner. They come from: 1) uniqueness of the digits in all the (i, j) Sudoku cells, 2) a digit must occur once and only once in each row, column and in each of the nine 3×3 subgrids, and 3) obeying the clues. Constraint type 1) was already expressed above, namely that for every cell (i, j) , in the set $(x_{ij}^1, \dots, x_{ij}^9)$ one and only one variable is TRUE, all others must be FALSE. Type 2) constraints are similar, e.g., in row i and layer L_a the set $(x_{i1}^a, \dots, x_{i9}^a)$ must contain one and only one TRUE variable, all others must be false and this must hold for all rows and layers, etc. Observe that all constraints are in the form of a set of 9 boolean variables of which we demand that one and only one of them be TRUE, all others FALSE. When this is satisfied, we say that the constraint itself (or “clause”) is satisfied, or TRUE. Such CSPs are called +1-in- k -SAT and they are part of so-called “locked occupation problems”, which is a class of exceptionally hard CSPs^{18,19}. Type 3) constraints are generated by the clues (or givens) which are symbols already filled in some of the cells and their number and positioning determines the difficulty of the puzzle. They are also set in a way to guarantee a unique solution to the whole puzzle. If there are given d clues, then this implies setting d boolean variables to TRUE, which means eliminating exactly $4d$ constraints of type 1) and 2) (one vertical or uniqueness constraint, one row, one column and one 3×3 subgrid



constraint). Thus, Sudoku is a +1-in-9-SAT type CSP with N boolean variables and $324 - 4d$ constraints of +1-in- k -SAT type ($k \leq 9$). N is a complicated function of the positioning of the clues. The example in Fig. 1a has $d = 22$ clues with $N = 232$ unknown boolean variables. In layer L_4 as illustrated in Fig. 1c there are 28 unknown boolean variables (white cells). These 28 variables appear in a total of 17 constraints of +1-in- k -SAT type. More precisely there is one +1-in-2-SAT, six +1-in-4-SAT, four +1-in-5-SAT and six +1-in-6-SAT type of constraints related to L_4 . The other layers generate the remaining $324 - 4d - 17 = 219$ of +1-in- k -SAT type constraints (with $k \leq 9$).

Since our continuous-time SAT solver has been designed to solve boolean satisfiability problems in conjunctive normal form (CNF), we need to bring the +1-in- k -SAT type problems above into this form, and thus formulate it as a k -SAT CNF problem. The CNF is a conjunction (AND, denoted by \wedge) of clauses each clause expressed as the disjunction (OR, denoted by \vee) of literals. For k -SAT in CNF there are N boolean variables $x_i \in \{0, 1\}$ and an instance is given as a propositional formula \mathcal{F} , which is the conjunction of M clauses C_m : $\mathcal{F} = C_1 \wedge \dots \wedge C_m \wedge \dots \wedge C_M$, with $C_m = z_{m1} \vee \dots \vee z_{mk_m}$, $k_m \leq k$ and z_j is a literal, that is $z_j \in \{x_j, \bar{x}_j\}$ (\bar{x}_j is the negation of x_j). According to a well known theorem of propositional calculus, all boolean propositions can be formulated in CNF using De Morgan's laws and the distributive law, and thus any +1-in- k -SAT type constraint as well. The Methods section describes how to translate +1-in- k -SAT type constraints into CNF.

Once the transformation to CNF is completed we are left with N variables and M clauses of the type described above, called CNF clauses from here on. In our case the number of variables appearing in a CNF constraint has the property $1 \leq k_m \leq 9$. The parameters N , M and $\{k_m\}_{m=1}^M$ depend on the clues that are difficult to express analytically, but easy to determine computationally. The puzzle from Fig. 1a is ultimately formulated as a CNF SAT problem with $N = 232$ variables and $M = 1718$ CNF clauses. An often used parameter of a satisfiability problem is the number of CNF constraints per variable, or *constraint density*, $\alpha = M/N$, also used as a typical hardness indicator, however, as we show below, this is not an accurate measure of hardness.

The continuous-time deterministic k -SAT solver. In Ref [15] a continuous-time deterministic solver was introduced to solve k -SAT problems in conjunctive normal form. The set of clauses specifying the constraints are translated into an $M \times N$ matrix: $C = \{c_{mi}\}$ with $c_{mi} = 1$ if the variable x_i is present in clause m in direct (non-negated) form, namely $x_i \in C_m$, $c_{mi} = -1$ if $\bar{x}_i \in C_m$ and $c_{mi} = 0$ if x_i and \bar{x}_i are both absent from C_m . To every variable x_i one associates a continuous spin variable $s_i \in [-1, 1]$ such that when $s_i = \pm 1$ then $x_i = (1 + s_i)/2 \in \{0, 1\}$, and to every clause C_m one associates the function:

$$K_m(s) = 2^{-k_m} \prod_{j=1}^{k_m} (1 - c_{mj}s_j), \quad m \in \{1, \dots, M\}. \quad (1)$$

We have $K_m \in [0, 1]$ for all $s \in [-1, 1]^N$. It is easy to check that $K_m = 0$ only for those $s_i \in \{-1, +1\}$ values for which the corresponding x_i -s satisfy clause C_m (otherwise we always have $K_m > 0$). That is, K_m plays the role of an energy function for clause C_m and its ground state value of $K_m = 0$ is reached if C_m is TRUE, and only then. We also need the quantities $K_{mi} = K_m/(1 - c_{mi}s_i)$ that is, with the i -th term missing from the product in (1). Clearly, $K_{mi} \in [0, 1/2]$. The continuous time dynamical system introduced in¹⁵ is defined via the set of $(N + M)$ ordinary differential equations (ODEs):

$$\frac{ds_i}{dt} = \sum_{m=1}^M 2a_m c_{mi} K_{mi}(s) K_m(s), \quad i = 1, \dots, N \quad (2)$$

$$\frac{da_m}{dt} = a_m K_m(s), \quad m = 1, \dots, M, \quad (3)$$

with the only requirements that $s_i(0) \in [-1, 1]$, $\forall i$ and $a_m(0) > 0$, $\forall m$. The latter implies from (3) that $a_m(t) > 0$, $\forall m, t$. It was shown in Ref [15] that system (2–3) always finds the solutions to k -SAT problems (encoded via the C matrix), when they exist, from almost all initial conditions (the exception being a set of Lebesgue measure zero). Here we give an intuitive picture for why that is the case. Due to (3) the auxiliary variables a_m grow exponentially at rate K_m . That is, the further is K_m from its ground-state value of 0, the faster a_m grows (in that instant). Moreover, the longer has K_m been away from zero, the larger is a_m , as seen from the formal solution to (3): $a_m(t) = a_m(0) \exp\left(\int_0^t d\tau K_m\right)$. Equation (2) can equivalently be written as a gradient descent on an energy landscape $V(s, a)$, that is $ds/dt = -\nabla_s V$, where ∇_s is the gradient operator in the spin variables and $V(s, a) = \sum_m a_m K_m^2(s)$. Clearly, $V \geq 0 \forall t$ and $V = 0$ if and only if s is a k -SAT solution, i.e., satisfies all the clauses ($K_m(s) = 0, \forall m$).

From the behavior of the a_m variables discussed above it also follows that the least satisfied constraints will dominate V (terms with the largest a_m -s). Without restricting generality, let the $a_1 K_1^2$ term be the most dominant at t . Then keeping only the dominant term on the rhs of (2) for those i for which $c_{1i} \neq 0$ we get $ds_i/dt = 2a_1 c_{1i} (1 - c_{1i}s_i)(K_1)^2$ or, equivalently: $d(1 - c_{1i}s_i)/dt = -(1 - c_{1i}s_i) 2a_1 (K_1)^2$. This shows that the term $(1 - c_{1i}s_i)$ is driven exponentially fast towards zero, that is towards satisfying K_1 (and all the other constraints containing this term). As K_1 decreases, some other constraint becomes dominant, and thus, in a continuous fashion, all constraints are driven towards satisfiability. The exponential growth guarantees that the trajectory is always pulled out of any potential well. When the problem is unsatisfiable, the system generates a chaotic dynamics in $[-1, 1]^N$, indefinitely. For more details about the properties of the continuous-time dynamical system (CTDS) (2–3) see Ref. [15].

Puzzle hardness as transient chaotic dynamics. Since Sudoku puzzles always have a solution, the corresponding boolean SAT CNF formulation also has a solution, and system (2–3) will always find it. The nature of the dynamics, however, will depend on the hardness of the puzzle as we describe next.

In Fig. 2a we show an easy puzzle with 34 clues (black numbers)²⁰. After transforming this problem into SAT CNF, we obtain $N = 126$ and $M = 717$, with a constraint density of $\alpha = M/N = 5.69$. As described above, in our implementation there is a continuous spin variable s_{ij}^a associated to every boolean variable x_{ij}^a in every 3D cell (i, j, a) . In the right panels of Fig. 2 we show the dynamics of the spin variables in the cells of the 3×3 grid formed by rows 4–6 and columns 7–9. The $s_{ij}^a(t)$ curves are colored by the digit a they represent ($a = 1, \dots, 9$) as indicated in the color legend of Fig. 2. The dynamics was started from a random initial condition. Indeed, our solver finds the solution very quickly, in about 15 time units, for the easy puzzle in Fig. 2a.

In Fig. 2b we show the dynamical evolution of variables for an ultra-hard Sudoku instance with only 21 clues. This puzzle has been listed as one of the world's hardest Sudokus, and even has a special name: “Platinum Blonde”^{21–23}, and it was the most “difficult” for our solver among all the puzzles we tried. After transforming it into SAT CNF, we obtain $N = 257$ variables and $M = 2085$ constraints. Not only that we have twice as many unknown variables but the constraint density $\alpha = M/N = 8.11$ is also larger than in the previous case, signaling the hardness of the corresponding SAT instance. The complexity of the dynamics in this case is seen in the right panel of Fig. 2b, exhibiting long chaotic transients before the solution is found at around $t \approx 150$. For an animation of the dynamics for a similarly hard puzzle¹⁴ see Ref [24].

We can also observe from the right panels in Fig 2 that there is one dominating digit (a -value), corresponding to which vertical cell at that given (i, j) grid cell has the largest s_{ij}^a value. This can be taken as the digit D_{ij} the solver is considering in the given grid cell (i, j) at that

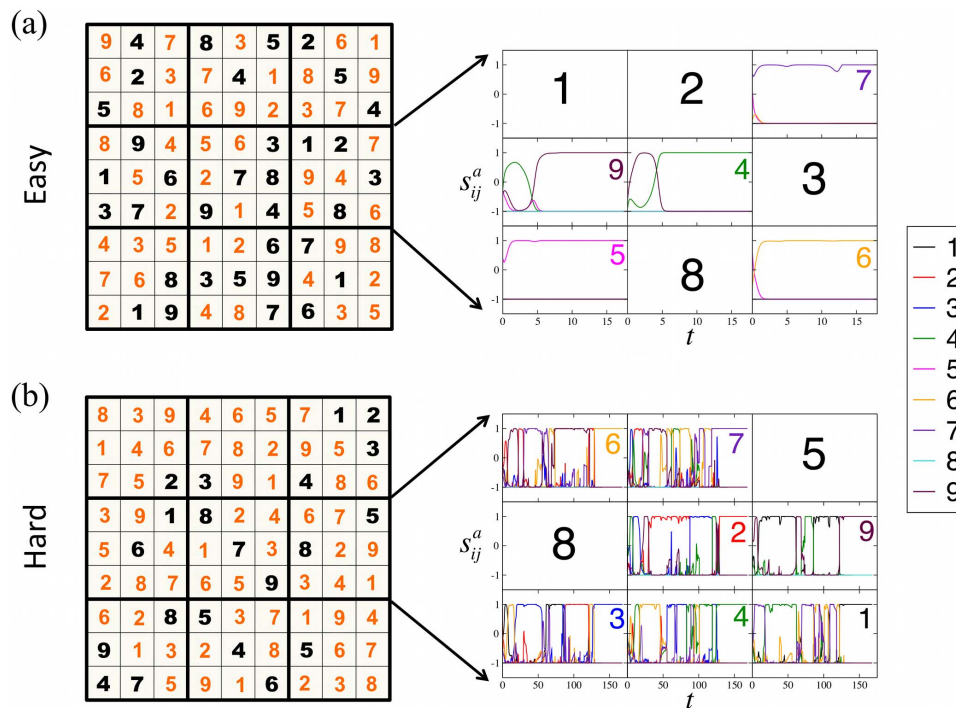


Figure 2 | Solving Sudoku puzzles with the deterministic continuous-time solver (2–3). (a) presents an easy puzzle with the evolution of the continuous-time dynamics shown within a 3×3 grid (rows 4–6, columns 7–9). (b) shows the same, but for a known, ultra-hard puzzle called “Platinum Blonde”^{21–23}. The trajectories in the right panels show the evolution of the analog variables $s_{ij}^a(t)$ colored by the corresponding digit a . Thus for each cell (i, j) we have 9 such running trajectories, but they cannot always be discerned as many of them are running on top of each other, close to -1, as seen in (a).

moment. We will use this observation to provide below an alternate illustration of the dynamics’ transiently chaotic behavior. Let us fix a random initial condition except for two chosen variables that are varied along the points of a square grid within the domain $[-1, 1]^2$. There is no particular relevance as to which pairs of variables are chosen to be varied, let us denote them by s_1 and s_2 . Let us choose an arbitrary empty cell (p, q) in the original Sudoku puzzle and monitor the dominating digit in it at time t . We will color the initial conditions in the plane (s_1, s_2) according to the dominating digit in (p, q) at time t . This will provide a map expressing the “sensitivity to initial conditions” that varies across time. Since all puzzles have solutions, the maps eventually assume one solid color according to the digit of the solution in the monitored cell, however, for hard puzzles, it may assume highly complex patterns before it does that, as shown in Fig. 3. Here we show these colormaps for the easy and hard Sudoku puzzles from Fig. 2 at times $t = 10, 15, 20$. For the easy puzzle (top row) the cell was chosen to be $(p, q) = (1, 1)$. At time $t = 10$ the whole map shows $D_{1,1} = 6$ (orange), which is not the solution digit (it is still searching for the solution). At time $t = 15$, however, we see two clearly separated domains, in one of them $D_{1,1} = 6$, in the other $D_{1,1} = 9$ (light blue), the latter being the correct digit. As time passes, the orange (incorrect) domain shrinks, because trajectories from an increasing number of initial conditions find the solution. At $t = 20$ almost the whole map shows the correct digit $D_{1,1} = 9$, except for a thin line.

In the case of the hard Sudoku puzzle (bottom row in Fig. 3, $(p, q) = (6, 8)$) more colors enter the picture over time, in a fractal-like pattern. Eventually the whole map becomes dark blue (digit 4) corresponding to the solution in this cell. Changing the initial condition slightly about this fractal set (which is really a time-dependent fractal boundary) may result in a completely different digit (color) being considered in cell (p, q) at time t . This sensitivity to initial conditions is indicative of the chaotic behavior of the (deterministic) search dynamics.

The appearance of transient chaos is a fundamental feature of the search dynamics and can be used to separate problems by their

hardness. In Ref [15] we have shown that within the *thermodynamic limit* ($N \rightarrow \infty, M \rightarrow \infty, \alpha = M/N = \text{const.}$) of random k -SAT ensembles this appears as a phase transition at the so-called chaotic transition point α_c in terms of the constraint density $\alpha = M/N$. Since there is no “thermodynamic limit” for 9×9 Sudoku problems ($N < 729$), one cannot define a simple order-parameter and use it to rate problem hardness in the same way¹⁵. However, once a problem is given, the corresponding dynamical system (2–3) is well defined, and so is its dynamical behavior. Even though we do not have a well-defined ensemble-based statistical order parameter (which has little meaning for specific SAT instances anyway), here we show next how can we use a well-known invariant quantity from non-linear dynamical system’s theory to categorize problem hardness for specific instances.

A Richter-type scale for Sudoku hardness. As suggested by the two examples in Fig. 2, the hardness of Sudoku puzzles correlates with the length of chaotic transients. A consistent way to characterize these chaotic transients is to plot the distribution of their lifetime. Starting trajectories from many random initial conditions, let $p(t)$ indicate the probability that the dynamics has not found the solution by time t . A characteristic property of transient chaos^{11,12} in hyperbolic dynamical systems is that $p(t)$ shows an asymptotic exponential decay: $p(t) \sim e^{-\kappa t}$, where κ is called the escape rate. For an easy to read text on transient chaos see the book¹³. The escape rate, a measurable quantity, theoretically can be expressed as a zero of the spectral determinant of the evolution operator corresponding to the dynamical system (2–3) and well approximated using the machinery of cycle expansions based on dynamical zeta functions¹². It is an invariant measure of the dynamics in the sense that it characterizes solely the chaotic nonattracting set in the phase space of the system, and it does not depend on the distribution of the initial conditions, its support, or the details of the region from where the escape is measured (as long as it contains the non-attracting set)¹¹.

In Fig. 4a we plot the distribution $p(t)$ in log-linear scale for several puzzles gathered from the literature. The distributions were obtained

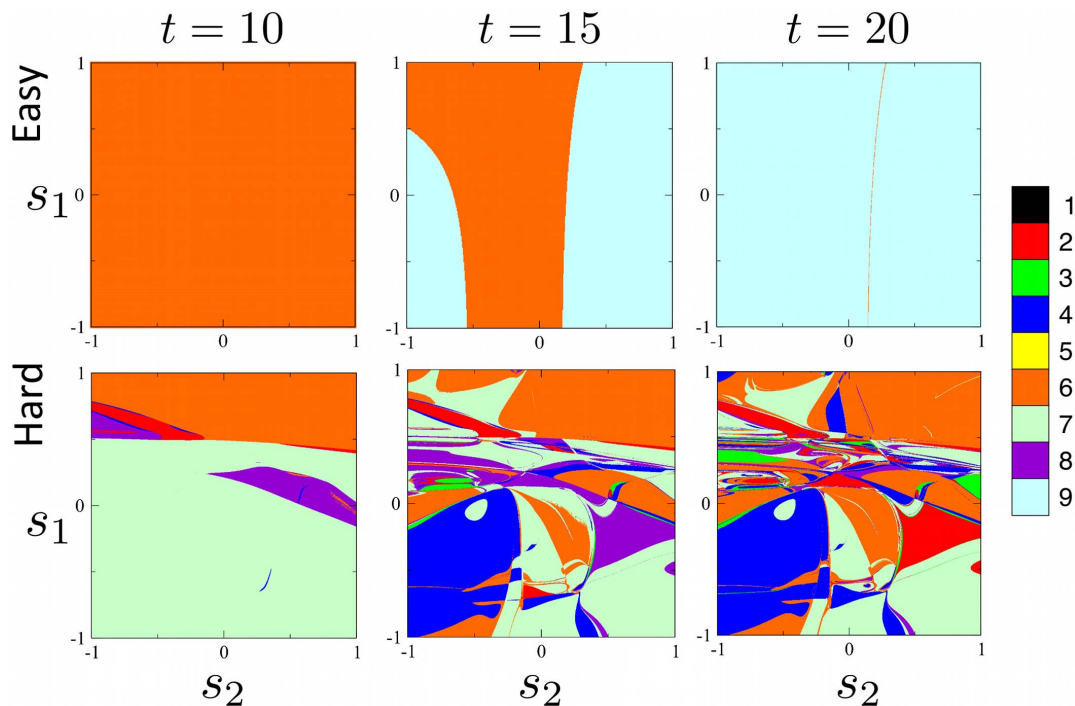


Figure 3 | Puzzle hardness as chaotic dynamics. We color the points of a $10^3 \times 10^3$ grid in an arbitrary plane (s_1, s_2) at time instant t according to the digit D_{pq} the solver is considering in an arbitrary but fixed cell (p, q) at that instant, given that we started the trajectory of the CTDS from those grid-points. For these initial conditions only the points in the (s_1, s_2) plane were varied, all other spin values were kept fixed at the same randomly chosen values. For the same easy problem as in Fig. 2, and for $(p, q) = (1, 1)$ (top row of panels) almost all initial conditions in this plane involve only two digits, and after $t = 20$ the corresponding trajectories have converged to the solution digit (9, light blue), except for a thin line, which, however, will also become light blue. The bottom row of panels shows the same for the hard problem of Fig. 2 based on what happens in the cell $(p, q) = (6, 8)$. The strong sensitivity to initial conditions appears as fractal structures of increasing complexity as time goes on, before eventually everything converges to the same color/digit (dark blue, corresponding to digit 4, not shown).

from over 10^4 random initial conditions. The decay shows a wide range of variation between the puzzles. For easy puzzles the transients are very short, $p(t)$ decays fast resulting in large escape rates but for hard puzzles κ can be very small. Fig. 4b shows a zoom onto the $p(t)$ of hard puzzles. For example, for the puzzle in Fig. 2a we obtain $\kappa = 0.156$, whereas for Fig. 2b (Platinum Blonde) the escape rate is $\kappa = 0.00026$. In spite of the large variability of the decay rates, we see that in all cases the escape is exponentially fast (the faster than exponential appearance for some of the curves in Figures 4a,b is due to finite size statistics and the finiteness of the time interval considered).

The several orders of magnitude variability of κ naturally suggests the use of a logarithmic measure of κ for puzzle hardness, see Fig. 4c, which shows the escape rates on a semilog scale as function of the number of clues, d . Thus, the escape rate can be used to define a kind of “Richter”-type scale for Sudoku hardness:

$$\eta = -\log_{10}(\kappa) \quad (4)$$

with easy puzzles falling in the range $0 < \eta \leq 1$, medium ones in $1 < \eta \leq 2$, hard ones in $2 < \eta \leq 3$ and for ultra-hard puzzles $\eta > 3$.

We chose several instances from the “Sudoku of the Day” website²⁰ in four of the categories defined there: easy (black square), medium (red circle), hard (green \times) and absurd (blue star). These ratings on the website try to estimate the hardness of puzzles when solved by humans. These ratings correlate very well with our hardness measure η , giving an average hardness value of $\langle \eta \rangle = 0.816$ for easy, $\langle \eta \rangle = 1.439$ medium, $\langle \eta \rangle = 1.782$ for hard and $\langle \eta \rangle = 1.809$ for what they call absurd. Another site we analyzed puzzles from is “Extreme Sudoku”²⁵ (brown + signs on Fig. 4). It claims to offer extremely hard Sudoku puzzles, their categories being: evil, excessive, egregious, excruciating and extreme. Indeed those puzzles are

difficult with a range of $\eta \in [1.1, 1.9]$ on the hardness scale, however, still far from the hardest puzzles we have found in the literature. Occasionally, daily newspapers present puzzles claimed to be the hardest Sudoku puzzles of the year. In particular, the escape rate for the Caveman Circus 2009 winner²⁶ (turquoise diamond) and the Guardian 2010 hardest puzzle²⁷ (maroon diamond) are indeed one order of magnitude smaller than the hardest puzzles on the daily Sudoku websites, placing them at $\eta = 2.93$ and $\eta = 2.82$ on the hardness scale. The USA Today 2006 hardest puzzle²⁸, however, does not seem to be that hard for our algorithm having $\eta = 2.17$ (magenta diamond). Eppstein²⁹ gives two Sudoku examples (orange left-pointing triangles) while describing his algorithm, one with $\eta = 1.288$ and a much harder one with $\eta = 2.017$. Elser *et al.*¹⁴ present an extremely hard Sudoku (black filled circle), which has an escape rate of $\kappa = 0.0023$ resulting in $\eta = 2.639$.

The smallest escape rates we have found are for the Sudokus listed as the hardest on Wikipedia^{22,23} (red triangles). The five puzzles, which we tested are called Platinum Blonde, Golden Nugget, Red Dwarf, coly013 and tarx0134. They have a hardness in the range $3 < \eta < 3.6$, the ultra-hard Platinum Blonde (shown in Fig. 2b) being the hardest with $\eta = 3.5789$ ($\kappa = 0.00026$).

While the escape rate correlates surprisingly well with human ratings of Sudoku hardness, it is natural to expect a correlation with the number of clues, d . Indeed, as a general rule of thumb, the fewer clues are given, the harder the puzzle, however, this is not universally true¹. Here we tested a few instances with minimal³⁰, that is 17 clues and almost minimal 18 clues (orange filled circles)^{31–33}. As seen from Fig. 4c, these are actually easier ($1.2 < \eta < 2.4$) than the hardest instances with more $d = 21, 22$ clues. In Fig. 4d we then plot the escape rate as function of the constraint density $\alpha = M/N$, leading to practically the same conclusion. This is because the constraint

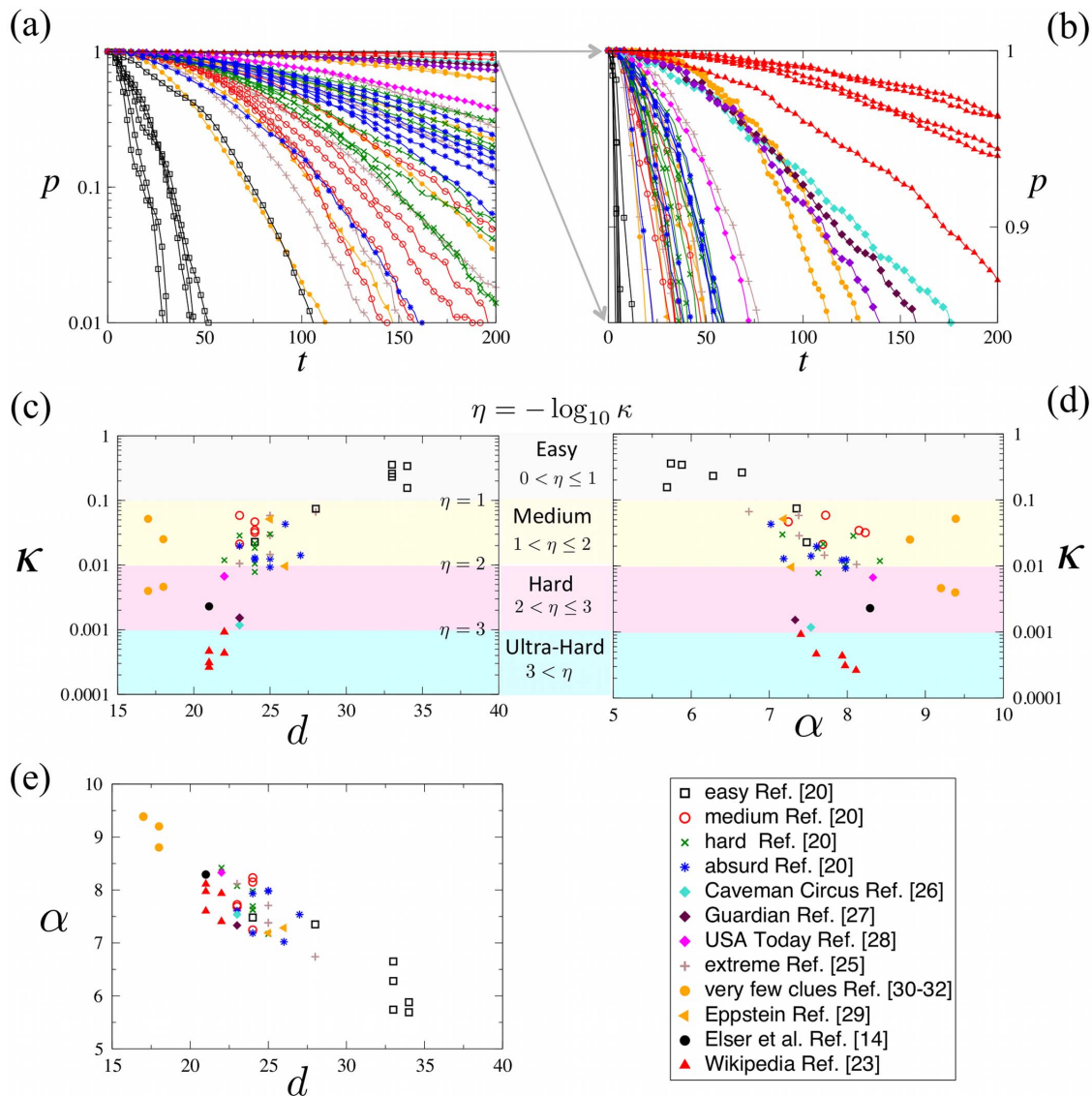


Figure 4 | Escape rate as hardness indicator. (a) shows the distribution in log-linear scale of the fraction $p(t)$ of 10^4 randomly started trajectories of (2–3) that have not yet found a solution by analog time t for a number of Sudoku puzzles taken from the literature (see legend and text) with a wide range of human difficulty ratings. The escape rate is obtained from the best fit to the tail of the distributions. (b) is a magnification of (a) for hard puzzles. (c) and (d) show the escape rate κ in semilog scale vs the number of clues d and constraint density α indicating good correlations with human ratings (color bands). (e) shows the relationship between the number of clues d and α for the puzzles considered.

density α is essentially linearly correlated with the number of givens d , as shown in Fig. 4e. The apparent non-monotonic behavior of puzzle hardness with the number of givens or constraint density is due to the fact that hardness cannot simply be characterized by a global, static variable such as d or α , but it also depends on the *positioning* pattern of the clues¹.

Discussion

Using the world of Sudoku puzzles, here we have presented further evidence that optimization hardness translates into complex dynamical behavior by an algorithm searching for solutions in an optimal fashion. Namely, there seems to be a trade-off between algorithmic performance and the complexity of the algorithm and/or its behavior. Simple, sequential search algorithms have a trivial description and simple dynamics, but an abysmal worst-case performance (2^N), whereas algorithms that are among the best performers are complex in their description (instruction-list) and/or behavior (dynamics). This happens because in order to improve performance, algorithms

have to exploit the structure of the problem one way or another. As hard problems have complex structures, the dynamics of the algorithms should be indicative of the problem's hardness. However, as a word of caution, observing complex dynamics performed by some black-box algorithm does not necessarily imply problem hardness. For example, one could consider any dynamical system that is guaranteed to visit all the 2^N discrete states/configurations but perhaps with some arbitrary, complex behavior. This algorithm will always find solutions if they exist (by checking for satisfiability after every new state visited). But its instruction list would have no relevance to the problem itself (apart from the checking instructions to see if the new state satisfies the problem) and thus, it could take long times to find solutions even for problems that are otherwise easily solved by other algorithms. Hence, dynamical properties can only be regarded as descriptors of problem hardness if they are generated by algorithms that: 1) exploit the structure of the state space of the problem and 2) they show similar or better performance compared to other algorithms on the same problems.



The continuous-time dynamical system¹⁵ (2–3) as a *deterministic algorithm* does have these features: 1) the search happens on an energy landscape $V = \sum_m a_m K_m^2$ that incorporates simultaneously *all* the constraints (problem structure) 2) it solves easy problems efficiently (polynomial time, both analog and discrete) and 3) it guarantees to find solutions to hard problems even for solvable cases where many other algorithms fail. Although it is not a polynomial cost algorithm, it seems to find solutions in continuous-time t that scales polynomially with N^{15} . These features and the fact that the algorithm is formulated as a deterministic dynamical system with continuous variables, allows us to apply the theory of nonlinear dynamical systems on CTDS (2–3) to characterize the hardness of boolean satisfiability problems. In particular, via the measurable escape rate κ , or its negative log-value η , we can provide a single-scalar measure of hardness, well defined for any *finite instance*. We have illustrated this here on Sudoku puzzles, but the analysis can be repeated on any other ensemble from NP. Having a mathematically well-defined number to characterize optimization hardness for specific problems in NP provides more information than the polynomial/exponential-time solvability classification, or knowing what the constraint density $\alpha = M/N$ is (the latter being a non-dynamic/static measure). Moreover, within the framework of CTDS (2–3), dynamical systems and chaos theory methods can now be brought forth to help develop a novel understanding of optimization hardness.

Methods

Type 1) constraints (main text) impose the uniqueness of the symbol D_{ij} in a given cell, expressed as a +1-in-9-SAT constraint on the vector:

$$(x_{ij}^1, x_{ij}^2, \dots, x_{ij}^9). \quad (5)$$

that is, *one and only one* component in the boolean vector above can be TRUE (or 1), all others must be FALSE (or 0). Having 9×9 cells in the puzzle, this gives in total 81, +1-in-9-SAT constraints.

Type 2) constraints on rows, columns and sub-grids further impose that in every layer L_a we have the following 27, +1-in-9-SAT constraints:

$$\text{Rows :} \quad (x_{i1}^a, x_{i2}^a, \dots, x_{i9}^a), \quad i = 1, \dots, 9 \quad (6)$$

$$\text{Columns :} \quad (x_{1j}^a, x_{2j}^a, \dots, x_{9j}^a) \quad j = 1, \dots, 9 \quad (7)$$

$$\text{Subgrids :} \quad \left(x_{m+1,n+1}^a, x_{m+1,n+2}^a, x_{m+1,n+3}^a, \right. \\ x_{m+2,n+1}^a, x_{m+2,n+2}^a, x_{m+2,n+3}^a, \\ \left. x_{m+3,n+1}^a, x_{m+3,n+2}^a, x_{m+3,n+3}^a \right), \quad m, n = 0, 3, 6 \quad (8)$$

with the same meaning as before, that is in all the boolean vectors above one and only one variable can be TRUE all others must be FALSE. Together with the 81 constraints of type 1) we thus have a total of $9 \times 27 + 81 = 324$ constraints in +1-in-9-SAT form.

Finally, type 3) constraints are imposed via d given digits or clues. It was only recently shown that uniqueness of a solution demands that $d \geq 17^{30}$. As discussed in the main text, each clue will eliminate 4 constraints: in its vertical tower, its column, its row and the 3×3 sub-grid containing the clue. For example, let us examine layer L_4 (Fig. 1c) of the puzzle shown in Fig. 1a. There are three clues of 4 in cells (1, 9), (3, 3), (4, 4) and thus $x_{1,9}^4 = 1, x_{3,3}^4 = 1, x_{4,4}^4 = 1$ have to be fixed as TRUE in L_4 . In order to satisfy the constraints, the other variables in the same rows, columns, blocks and vertical columns must be set to FALSE. The unknown variables left in the SAT problem will be those in the light cells of Fig. 1c. (The other clues will eliminate constraints and variables in other layers and vertical columns.) The total number of unknown variables N depends on d and on the placement of clues. The number of constraints is always $324 - 4d$, however the number of variables in a clause can vary. For example in Fig. 1c the constraint corresponding to the second row in layer L_4 has only 2 unknown variables left (+1-in-2-SAT).

After the unknown boolean variables and the constraints have been identified we need to transform the formula into CNF. This is done easily, as any +1-in- k -SAT

clause defined on the (y_1, y_2, \dots, y_k) variables can be written as one k -SAT and $k(k-1)/2$ of 2-SAT constraints:

$$(y_1 \vee y_2 \vee \dots \vee y_k) \wedge \left[\bigwedge_{i < j} (\bar{y}_i \vee \bar{y}_j) \right] \quad (9)$$

The disjunction (\vee) of the first k variables enforces that at least one variable must be true, but the rest of $(k(k-1)/2)$ 2-SAT type constraints ensure that only one of them is allowed to be true.

- Rosenhouse, J. & Taalman, L. *Taking Sudoku Seriously: The Math Behind the World's most Popular Pencil Puzzle* (Oxford University Press, New York, 2011).
- Garey, M. R. & Johnson, D. S. *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W. H. Freeman & Co., New York, NY, USA, 1990).
- Karp, R. M. Reducibility among combinatorial problems. In *Complexity of Computer Computations.*, R.E. Miller and J.W. Thatcher (editors). *Proc. of a Symp. on the Complexity of Computer Computations.* (New York: Plenum. pp. 85–103, 1972).
- Yato, T. & Seta, T. Complexity and completeness of finding another solution and its application to puzzles. *IEICE Trans. Fundamentals* E86-A(5), 1052–1060 (2003).
- Fortnow, L. The status of the P versus NP problem. *Commun. ACM* **52**, 78–86 (2009).
- Barahona, F. On the computational complexity of Ising spin glass models. *J. Phys. A: Math. Gen.* **15**, 3241–3253 (1982).
- Istrail, S. Statistical Mechanics, Three-Dimensionality and NP-Completeness: I. Universality of Intractability of the Partition Functions of the Ising Model Across Non-Planar Lattices. *Proceedings of the 32nd ACM Symposium on the Theory of Computing (STOC00)*, ACM Press, pp. 87–96 (2000).
- Brueggemann, T. & Kern, W. An improved local search algorithm for 3-SAT. *Theor. Comp. Sci.* **329**(1–3), 303313 (2004).
- Ott, E. *Chaos in Dynamical Systems* 2nd edn (Cambridge Univ. Press, 2002).
- Cencini, M., Cecconi, F. & Vulpiani, A. *Chaos: from simple models to complex systems* (World Scientific, Singapore, 2009).
- Lai, Y.-C. & Tél, T. *Transient Chaos: Complex Dynamics on Finite-Time Scales* (Springer 2011).
- Cvitanović, P., Artuso, R., Mainieri, R. Tanner, G. & Vattay, G. *Chaos: Classical and Quantum*, ChaosBook.org/version13 (Niels Bohr Institute, Copenhagen 2010).
- Tél, T. & Gruiz, T. *Chaotic dynamics: and introduction.* (Cambridge University Press, 2006).
- Elser, V., Rankenburg, I. & Thibault, P. Searching with iterated maps. *Proc. Natl. Acad. Sci. USA* **104**, 418–423 (2007).
- Ercsey-Ravasz, M. & Toroczkai, Z. Optimization hardness as transient chaos in an analog approach to constraint satisfaction. *Nature Physics* **7**, 966–970 (2011).
- Kadanoff, L. P. & Tang, C. Escape from strange repellers. *Proc. Natl. Acad. Sci.* **81**, 1276–1279 (1984).
- Tél, T. & Lai, Y.-C. Chaotic transients in spatially extended systems. *Phys. Rep.* **460**, 245–275 (2008).
- Zdeborová, L. & Mézard, M. Locked constraint satisfaction problems. *Phys. Rev. Lett.* **101**, 078702 (2008).
- Zdeborová, L. & Mézard, M. Constraint satisfaction problems with isolated solutions are hard. *J. Stat. Mech.: Theor. Exp.* P12004 (2008).
- <http://www.sudokuoftheday.co.uk> (accessed on 26-07-2012)
- <http://forum.enjoysudoku.com/the-hardest-sudokus-new-thread-t6539.html> (accessed on 23-07-2012).
- http://wiki.karadimov.info/index.php/Sudoku_algorithms#Exceptionally_difficult_Sudokus_.28h (accessed on 25-07-2012).
- http://en.wikipedia.org/wiki/Algorithmics_of_sudoku#Exceptionally_difficult_Sudokus_.28hard (accessed on 27-07-2012).
- http://www.youtube.com/watch?v=y4_aSLP9g_w (accessed on 19-06-2012).
- <http://www.extremesudoku.info/sudoku.html> (accessed on 26-07-2012).
- <http://cavemancircus.com/2009/11/05/the-hardest-sudoku-puzzle-ever/> (accessed on 26-07-2012).
- Hutchinson, P. The world's hardest Sudoku? *The Guardian*, August 22 (2010) <http://www.guardian.co.uk/media/2010/aug/22/worlds-hardest-sudoku> (accessed on 21-07-2012).
- http://www.usatoday.com/news/offbeat/2006-11-06-sudoku_x.htm (accessed on 26-07-2012).
- Eppstein, D. Solving Single-digit Sudoku Subproblems, <http://arxiv.org/abs/1202.5074v2>
- McGuire, G., Tugeman, B. & Civario, G. There is no 16-Clue Sudoku: Solving the Sudoku Minimum Number of Clues Problem. <http://arxiv.org/abs/1201.0749> (accessed on 25-07-2012).
- <http://mapleta.maths.uwa.edu.au/gordon/sudokumin.php> (accessed on 26-07-2012).
- http://en.wikipedia.org/wiki/File:Symmetrical_18_clue_sudoku_01.JPG (accessed on 15-07-2012).



33. <http://www.flickr.com/photos/npcomplete/3603730706/> (accessed on 24-06-2012).

Acknowledgments

This work was supported in part by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, grant number PN-II-RU-TE-2011-3-0121 (MER) and by a University of Notre Dame internal capitalization grant (ZT).

Author contributions

The authors contributed analysis and tools equally. Both authors wrote and reviewed the manuscript.

Additional information

Competing financial interests The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

How to cite this article: Ercsey-Ravasz, M. & Toroczkai, Z. The Chaos Within Sudoku. *Sci. Rep.* **2**, 725; DOI:10.1038/srep00725 (2012).