npj | Computational Materials

www.nature.com/npjcompumats

ARTICLE OPEN

Autonomy in materials research: a case study in carbon nanotube growth

Pavel Nikolaev[1,2], Daylond Hooper[1,2,4], Frederick Webber[1,2,5], Rahul Rao[1,2], Kevin Decker[1,2], Michael Krein[3], Jason Poleski[3], Rick Barto[3] and Benji Maruyama[1]

Advances in materials are an important contributor to our technological progress, and yet the process of materials discovery and development itself is slow. Our current research process is human-centred, where human researchers design, conduct, analyse and interpret experiments, and then decide what to do next. We have built an Autonomous Research System (ARES)—an autonomous research robot capable of first-of-its-kind closed-loop iterative materials experimentation. ARES exploits advances in autonomous robotics, artificial intelligence, data sciences, and high-throughput and *in situ* techniques, and is able to design, execute and analyse its own experiments orders of magnitude faster than current research methods. We applied ARES to study the synthesis of single-walled carbon nanotubes, and show that it successfully learned to grow them at targeted growth rates. ARES has broad implications for the future roles of humans and autonomous research robots, and for human-machine partnering. We believe autonomous research robots like ARES constitute a disruptive advance in our ability to understand and develop complex materials at an unprecedented rate.

npj Computational Materials (2016) **2,** 16031; doi:10.1038/npjcompumats.2016.31; published online 21 October 2016

INTRODUCTION

Technological progress has been exponential in modern times: Powered flight progressed from the Wright Brothers' Flyer to the Apollo Moon Landing in less than 70 years, and Moore's Law recently celebrated its 50th anniversary of continued doubling of computing technology every 2 years. Advances in materials have enabled much of our technological progress,[1] but the process of materials development itself remains slow, often 20 years from invention to use.[2] One reason is that current methods of research are labour intensive. Humans drive most aspects of research, from initial hypothesis generation to experimental design, analysis, interpretation and iterative hypothesis refinement.

Many materials problems are complex and high dimensional. The number of experiments needed to capture the relevant phenomena is often overwhelmingly large and slows research. To address this, automation has been applied to research processes via high-throughput and combinatorial techniques in the fields of life sciences[3–7] and materials research.[8–16] In addition to automation, computation is being exploited to speed research. The Materials Genome Initiative[17] and related Integrated Computational Materials Science and Engineering efforts[18] have highlighted the need for computational approaches to more effectively make use of the explosion in data generation. However, neither automation nor Integrated Computational Materials Science and Engineering addresses the iterative nature of the research process. Each new set of experiments brings new knowledge to incorporate into the design of subsequent experiments. Here the human involvement in analysis and experimental design for high-throughput approaches remains a bottleneck.[3] The solution to these problems is to team human researchers with autonomous research robots that can capture and incorporate new experimental knowledge in an iterative learning loop, enabling us to traverse complex experimental parameter spaces.[19]

Autonomy is a topic of great current interest, and is an extension of automation; an autonomous system is capable of changing its behaviour in response to a changing environment. Traditional examples include driverless cars,[20] spacecraft[21] and unmanned aerial vehicles.[22,23] In autonomous systems, models of the environment are built up and decisions are exercised based on sensory data. In the context of scientific research, autonomy can be applied to systems where available instrumentation is employed as sensory input. Autonomy has already impacted the life sciences, exemplified by excellent work in the study of gene functions[24–26] and drug discovery.[27,28] However, autonomy has yet to deeply impact materials discovery and research. Here we demonstrate our Autonomous Research System (ARES), which we believe is the first of its kind to combine robotics, artificial intelligence (AI), data science and *in situ* techniques towards the vastly complex realm of materials development. In particular, we applied ARES towards the synthesis of single-walled carbon nanotubes (CNTs).

CNTs have unparalleled properties, combining strength and stiffness with extreme electrical and thermal conductivity. These properties are being exploited in nanoelectronic and plasmonic devices, lightweight conductive cables, composites and ballistic materials, as well as biological and chemical sensors.[29,30] But the synthesis of CNTs remains poorly controlled and poorly

[1]Air Force Research Laboratory, Materials and Manufacturing Directorate, RXAS, WPAFB, Dayton, OH, USA; [2]UES Inc, Dayton, OH, USA and [3]Lockheed Martin Advanced Technology Laboratories, Cherry Hill, NJ, USA.
Correspondence: B. Maruyama (benji.maruyama@us.af.mil) or P Nikolaev (pavel.nikolaev.ctr@us.af.mil)
[4]Present address: Air Force Research Laboratory, RHCI, WPAFB, OH 45433 and Infoscitex Inc, Dayton, OH 45431, USA.
[5]Present address: Air Force Research Laboratory, 711 Human Performance Wing, Human Effectiveness Directorate, RHAS, WPAFB, Dayton, OH 45433, USA.
Received 25 April 2016; revised 26 August 2016; accepted 2 September 2016

understood, despite two decades of intensive effort. The defective, impure material that generally results from typical growth methods also lacks the diameter and helicity control needed for electronics, and the poor yield and product variability increases cost and impedes transition of this important material. One of the biggest challenges to controlling growth is the overwhelmingly large experimental parameter space (e.g., temperature, pressure, gas composition). Here we report the successful demonstration of autonomy applied to the experimental study of CNT synthesis, wherein ARES autonomously traversed high-dimensional parameter space, varying multiple growth parameters between each experiment with the goal of learning to control CNT growth rate. This work paves the way for ARES to be expanded to other materials and processes, and signifies a fundamental change in the process of research by teaming human and robot researchers to attack vastly more complex and potentially impactful problems.

RESULTS

With ARES, we successfully built and demonstrated an autonomous research system capable of learning to grow CNTs by developing an AI planner and linking it to an automated growth reactor with *in situ* characterisation used as feedback control to iteratively define, conduct and analyse experiments. ARES continuously improved its ability to target growth rates over a series of hundreds of experiments. Each robotically controlled experiment consisted of synthesising CNTs by a chemical vapour deposition (CVD) process where the ARES AI planner supplied the growth conditions—temperature, pressure, and partial pressures of ethylene, hydrogen and water vapour.

A unique feature of ARES is the use of a laser to individually heat small (10 μm) pillars of silicon pre-seeded with catalyst[15,31–34] for the CVD reaction (see experimental details and Supplementary Figure S1 for a schematic diagram of the instrumentation). We chose cobalt as the growth catalyst (2 nm film on a 10 nm thick alumina support), which was deposited uniformly onto the micro-pillars. Each pillar was thermally isolated on a wafer of thousands of pillars, making it in effect an independent micro-reactor. This enabled hundreds of experiments to be conducted in series by moving from one pillar to the next, with the ability to change all the experimental input conditions (i.e., temperature, pressure and gas composition). In addition, because of our unique configuration, ARES conducted its series of experiments without the need to exchange substrates from the growth chamber, thus enabling such high experimental throughput with minimal human intervention. Recently we demonstrated rapid automated experimentation with rates of up to 100 experiments per day (compared to 1 per day for conventional methods) for a series of 534 experiments in multi-dimensional parameter space.[15]

The ARES heating laser also served as the excitation source for Raman spectroscopy, enabling *in situ* acquisition of spectra during CNT growth. The characteristic Raman peak from CNTs, called the G band, corresponds to tangential vibrations of the carbon atoms, and its intensity is representative of the yield.[31] ARES uses the increase in intensity of the G band with time during each experiment to determine the CNT growth rate, v_{exp} (taken as the maximum experimentally observed growth rate, see Supplementary Figure S2). The experimental growth rate was used as the signal for the ARES feedback loop.

Our overall objective for ARES was to autonomously learn to control the growth rate of CNTs using AI and closed-loop feedback over many experimental iterations. And so, building on our previous work, we implemented an AI planner that proposed new experimental growth conditions based on an analysis of a database of prior experiments, iteratively improving its ability to predict growth rates. The AI experimental planner was comprised of a random forest model[35] with growth conditions exercised through a genetic algorithm.[36] This combination was chosen for its ability to capture nonlinear relationships between input and output variables within disjoint design spaces,[37] which we deemed appropriate for the complexity of nanotube growth. It is currently implemented in a customised version of the Lockheed Martin Nanotechnology Material Data Mining, Modeling & Management (NMD-M3) software tool.[38] Before the first autonomous experiments, an initial set of 84 experiments was conducted to provide a database of prior knowledge needed for the random forest planner to build its first model. These experiments were designed to span the growth parameter input space in a grid style and were executed in automated mode, i.e., with pre-planned conditions supplied by the user but executed without user intervention.

Once the initial database was established, ARES performed a series of more than 600 experiments in autonomous mode, where the AI planner generated input conditions for each experiment. For each experimental iteration the AI planner received an objective growth rate from the user. It then analysed the database of prior experiments and generated new experimental growth conditions expected to achieve a predicted growth rate, v_{pred}, that targeted the user-supplied objective growth rate. Because the database incompletely spans the experimental parameter space, the predicted growth rate can differ from the objective, a known effect arising from domain applicability.[39] After each experiment the database of prior experimental results was updated with the resultant experimental growth rate, and the planner refined its model representation to reflect the latest information. Over the course of the experimental campaign the user periodically modified the objective growth rate to cause ARES to probe a broader span of the experimental parameter space, which increased its domain coverage. To test for convergence between the predicted and experimentally achieved growth rates, series of experiments were grouped into tasks of 29–94 experiments. The objective growth rate was held constant for later tasks, but varied within some earlier tasks (see Supplementary Table S1).

The results of the convergence test are shown in Figure 1a, which compares the experimental growth rates to the growth rates predicted by ARES. Note that as ARES gained more experience, i.e., as the cumulative number of experiments increased, the spread between experimental and predicted growth rates became smaller, thus demonstrating experimental convergence. We quantified convergence by normalising the difference between experimentally obtained and predicted growth rates: $\Delta = \frac{v_{exp} - v_{pred}}{v_{pred}}$ in Figure 1b. The mean values of Δ, μ_Δ and their s.d.s., σ_Δ, for each task enabled a statistical treatment of the data to analyse the trend towards convergence. As ARES learned to target growth rates, the mean difference from the prediction trended towards zero, implying that it successfully predicted the experimental growth rate. Moreover, the s.d. reduced to approximately 30%. To understand the context of these s.d. values, we conducted a series of experiments using the same input conditions, and analysed the statistical spread in the experimental growth rate. We found that the intrinsic variability in the system, which we termed the noise floor, over 20–30 experiments ranged from 20 to 30%, which is similar to the value found by Oliver *et al.* for their automated CNT growth furnace.[14] By the end of the experimental campaign the variability in the experimental values matched the noise floor. Thus we conclude that ARES was able to target growth rates to the degree of variability intrinsic to our system. Hereafter we use 'on-target' to refer to experiments whose growth rates matched the predicted growth rates within the variability of the system.

Having achieved experimental convergence we endeavoured to analyse the results with *ex situ* characterisation and data mining. We extracted a subset of experiments that achieved the targeted growth rate. Scanning electron microscopy imaging after growth confirmed that experimental growth rates measured *in situ* were

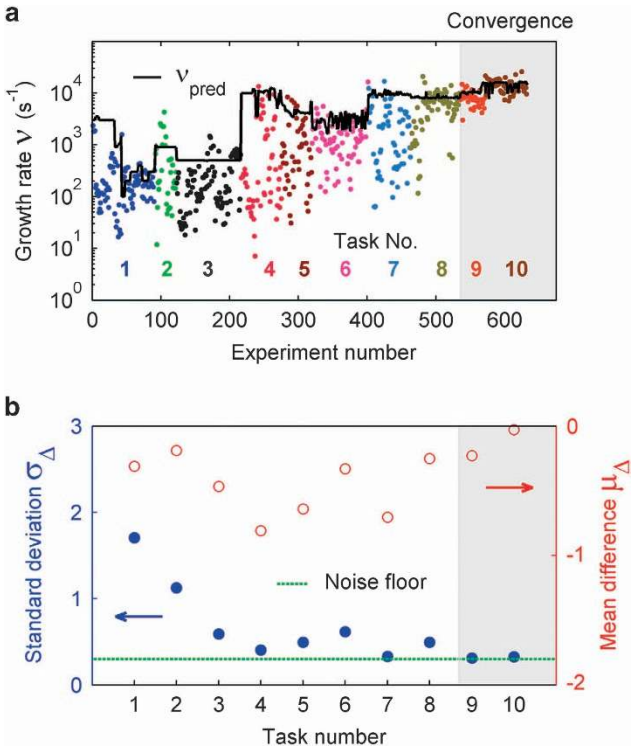

Figure 1. Demonstration of ARES learning to target CNT growth rates. (**a**) Experimental and predicted growth rates. (**b**) Mean (μ_Δ)and s.d. (σ_Δ) of the normalised difference between experimental and predicted growth rates. Initially the experimental rates were scattered far from the predicted (i.e., large σ_Δ and μ_Δ deviating from zero). As ARES learned over hundreds of experimental iterations the experimental rates converged (shaded area) to the predicted ones. This resulted in μ_Δ trending towards zero, and σ_Δ trending towards 30%, the noise floor of the system.

commensurate with CNT yield. Figure 2 shows the results of growth experiments that achieved on-target rates of 500, 3,000 and 16,000 s^{-1}, and demonstrates that the density of CNTs in the scanning electron microscope images is proportional to the experimentally observed growth rates.

We then analysed the progression of experimental conditions selected by ARES that led to growth rate convergence. In each iteration, the initial sampling of growth conditions from the genetic algorithm was inherently stochastic, yielding a ranked list of suggested experiments where predictions closest to the objective were prioritised. In order to avoid proposing the same experiment repeatedly, suggestions from using the genetic algorithm were further filtered based on proximity to existing data (assessed by Euclidian distance), to prioritise selection of 'different' experiments. In an early task, when models were trained from a limited and sparse data-set, this strategy led to emphasis on a broad range of parameter choices (Figures 3a and b) and resulted in few experiments matching the predictions: Only 8% were on-target. In a later task (Figures 3c and d) models were trained from a three times larger data-set. The resulting improvement in models' fidelity led to better match between predictions and experiments (68% of on-target experiments) and a narrow range of experimental parameters selected by the filtering algorithm. The narrow ranges of parameter choices in Task 10 are significant, representing convergence on a set of growth conditions that can be used as a growth recipe.

We then analysed the entire database of experiments for insights into growth kinetics. The average water/ethylene ratio in Task 10 is 1.6×10^{-3} (Supplementary Figure S3a), which is comparable to the 1×10^{-3} reported to maximise CNT yield and

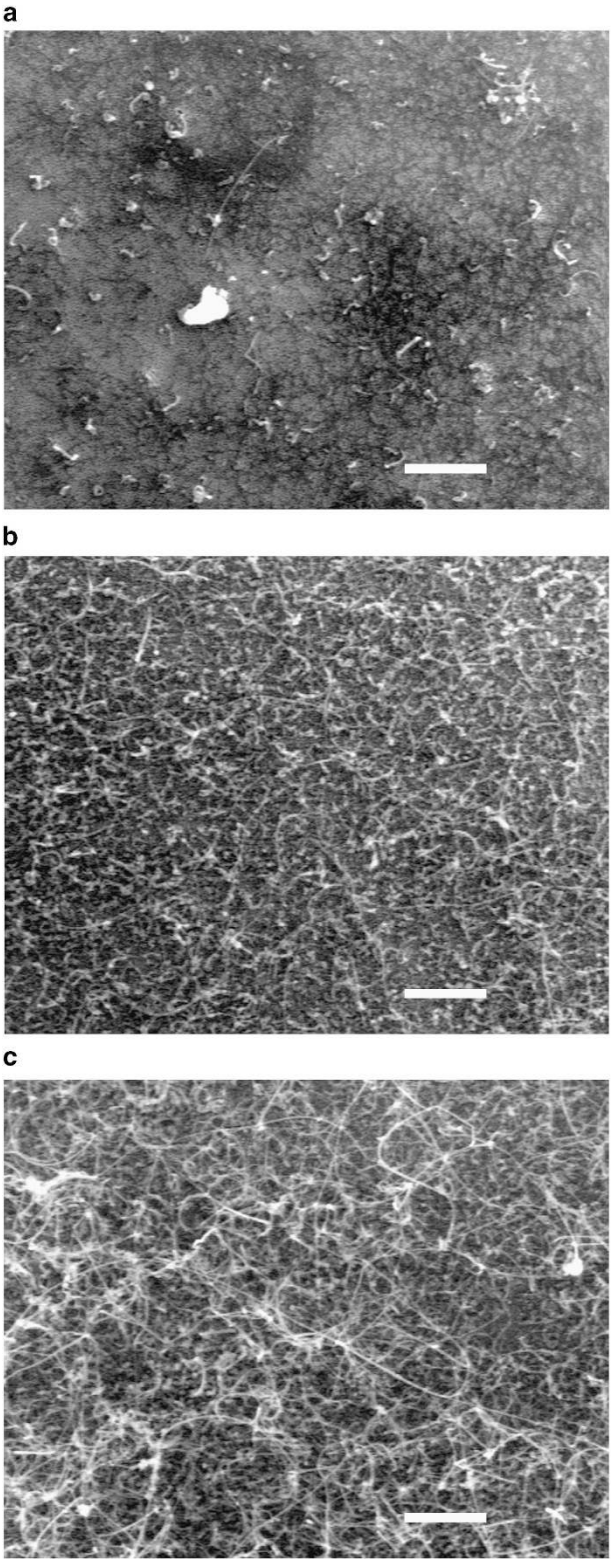

Figure 2. SEM images demonstrating correlation between growth rate and yield of CNTs. On-target experiments with growth rates near the predicted ones of (**a**) 500, (**b**) 3,000 and (**c**) 16,000 (s^{-1}). The time over which the nanotubes grew was approximately the same. The amount of nanotubes in the images increased in proportion to the growth rate. Scale bars: 500 nm.

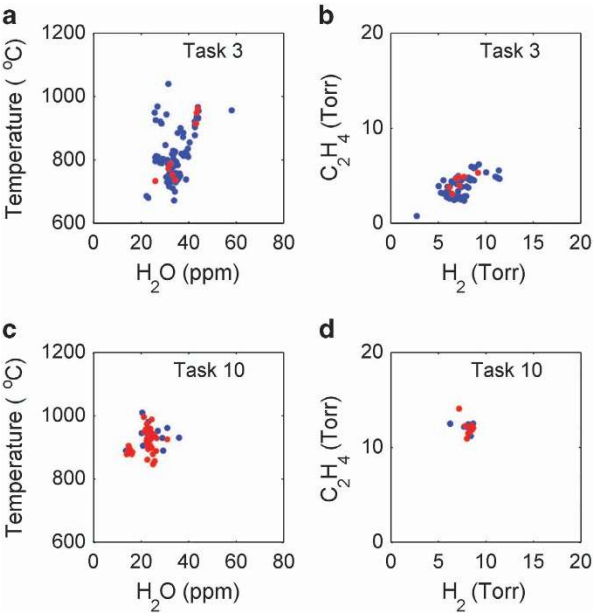

Figure 3. Variability in the experimental parameter space with learning. Experimental conditions chosen by ARES in Task 3, before convergence (**a**, **b**), and in Task 10, after convergence (**c**, **d**), are compared over four experimental parameters (temperature, water concentration, and H_2 and C_2H_4 partial pressures). Red dots represent successful, on-target experiments. (**a**, **b**) Before convergence ARES sampled a wide range of growth conditions, and only 8% of experiments were on-target. (**c**, **d**) After convergence ARES sampled a narrow range of growth conditions, with 68% on-target experiments, demonstrating its ability to autonomously optimise multiple experimental parameters.

growth rate by Futaba *et al.*[40,41] We also found that the rate dependence on the temperature (Supplementary Figure S3b) corresponds to the well-established Puretzky model[42] with activation energy $E_a = 1.1 \pm 0.3$ eV, which is within the 0.6–1.5 eV range reported in the literature and is interpreted as the activation energy of the precursor decomposition reaction on the catalyst surface.[42–44]

The data mining results highlight ARES' ability to search across a complex, multi-dimensional parameter space. All of these growth parameters are known to significantly affect CNT growth, and typically require extensive experimentation and serial optimisation by conventional methods.[40,41] The ability of the ARES planner to optimise multiple experimental parameters simultaneously and converge on the predictions highlights its efficacy in solving complex materials research problems that challenge human researchers using conventional, non-autonomous research processes.

DISCUSSION

We believe ARES is the first demonstration of an autonomous research system in the area of materials science. The ability of ARES to learn to control the growth rates of CNTs in multi-dimensional parameter space and over a range of rates is an inroad to the complex and vexing challenge of controlled nanotube synthesis. It also opens broad areas of applicability to other materials research problems, including other CVD reactions, additive manufacturing, metal oxidation, etc. In general, the advantage of the ARES method lies in the application of artificial intelligence and autonomy to research problems that are automatable, have large, complex experimental parameter spaces, and can be probed by *in situ* characterisation techniques to enable closed-loop feedback. These advantages come at the expense of the time needed for hardware and software development. However, as ARES development expands to more materials development problems, we expect the time and resources needed to adapt it to a particular research problem will progressively reduce. The iterative closed-loop approach contrasts with high-throughput/combinatorial techniques in important ways—they do not have autonomous feedback, and have slow iteration rates.[3] In addition, the broadcast nature of library generation results in large numbers of uninformative experiments.

Beyond faster and more efficient experimentation, ARES changes the size, complexity and risk of problems undertaken by researchers who operate under constrained resources. Because ARES enables more challenging experimental campaigns we expect ARES to be a disruptive tool for the research process itself. The immediate future for ARES includes the planner development to enable targeting of multiple objectives as well as expanding the diversity of AI planning algorithms. Random forest/genetic algorithm was just the first learning algorithm we used. ARES enables the application of a variety of AI learning approaches, with the distinct advantage of its iterative method driving the continuous updating and refinement of the AI planner. We will also incorporate more human-interpretable prior knowledge in the form of physical and chemical models, simulations and materials.

ARES also has implications for the role of human researchers and human-machine interaction, which is a topic of interest in the autonomous robotics community. In our case, humans chose the experimental area (CNT growth), the feedback signal and objective (maximum growth rate), seeded the database to initialise the AI planner, and periodically modified the objectives in response to the observed experimental progress. As we continue to develop autonomous research methods, ARES creates a need to explore and define human-machine interaction for generalised research processes, including goal definition, experimental path planning, progress towards goals, human-machine trust (e.g., awareness and reporting of the system and environmental state), and oversight. In a future where repetitive, mundane tasks are efficiently dispatched by ARES, and AI analysis of high-dimensional parameter spaces is done iteratively, researchers will be freed to work to their strengths, performing creative, insightful and contextual tasks.

MATERIALS AND METHODS

The ARES instrument diagram is shown in Supplementary Figure S1. CNTs were grown in a cold-wall CVD chamber installed on a two-axis motion stage above an inverted Raman microscope (Eclipse-Ti, Nikon, Mellville, NY, USA). The growth wafer consisted of 5×5 arrays of silicon pillars spaced 50 μm apart on a thermally insulating SiO_2 sublayer, which were 10 μm tall and 10 μm in diameter (Supplementary Figure S1 inset). The pillars were coated with a 10 nm ALD (atomic layer deposition) alumina support layer, on which a 2 nm Co catalyst film was deposited by electron beam evaporation. Each pillar constituted an independently addressable microreactor, which was heated by a 532 nm laser (Verdi V6, Spectra-Physics, Santa Clara, CA, USA) that doubled as the Raman excitation source through a 50× extra-long working distance objective lens. Raman spectra were collected from a ~5 μm size spot illuminated by the excitation laser. The small thermal mass of the pillar combined with the low thermal conductivity of the SiO_2 sublayer enabled heating to reaction temperatures within a fraction of a second when laser power was varied in the 0–2 W range. The CVD chamber pressure was measured by a capacitance pressure gauge (MKS) and regulated by a throttle valve (MKS) exhausting into a vacuum pump. H_2, C_2H_4 and Ar feedstock gases ware metered in varying ratios via mass flow controllers (MKS, Andover, MA, USA). Water vapour was introduced and controlled through a throttle valve (MKS), and its concentration was kept constant via a PID control loop with a dew point sensor (Shaw, Bradford, UK) as a feedback. The temperature-induced shift of the Si stokes and anti-stokes Raman bands (±520 cm^{-1}) was used to

calculate the growth temperature according to the following equation:

$$\Delta\omega(T) = C\left[1+\frac{2}{e^{\hbar\omega_0/2k_BT}-1}\right]+D\left[1+\frac{3}{e^{\hbar\omega_0/3k_BT}-1}+\frac{3}{\left(e^{\hbar\omega_0/3k_BT}-1\right)^2}\right] \quad (1)$$

Here k_B is Boltzmann's constant, $\hbar$ is Plank's constant, and ω_0, C and D are constants with the values 528, −2.96 and −0.174 cm^{-1}, respectively.[45] The accuracy of the temperature measurement was estimated to be within ~10–20 °C.

The two-axis motion stage, microscope focus (z), mass flow controllers, pressure controller, water controller, laser, shutters and spectrometer were all simultaneously controlled by custom software developed for 64-bit Windows 7 in C#/.NET 4.0 using Windows Visual Studio 2010. In a typical series of experiments, the planner was first given an objective. The database of the previous experiments containing inputs (temperature, pressure, gas flow rates, etc.) and outputs (kinetic parameters) was transferred to the planner. The planner converted flow rates and pressure into partial pressures of the constituent gases, built a model correlating inputs and outputs, and generated experimental conditions and predicted growth rate. The first pillar was positioned under the laser and the growth atmosphere was set to the intended composition. The laser power was increased in less than a second, and *in-situ* Raman spectra were acquired in 5 s intervals, with each spectrum processed in real time to obtain temperature (from Si bands' shifts) and CNT G- and D-band areas. Nanotube nucleation and growth was detected by the appearance and increase in intensity of G- and D-bands seen at ~1,590 and ~1,350 cm^{-1} (Supplementary Figure S1a). The G-band area was normalised by the laser power, by the Bose thermal occupation factor $[1-\exp(\hbar\omega/k_BT)]^{-1}$ (ω is G-band frequency, T is temperature), and by the area of the Si stokes peak at 520 cm^{-1} obtained with 200 mW laser power and 30 s exposure that served as a proxy for the spectrometer throughput. Normalisation was necessary in order to compare G-band areas and rates across different experiments, and resulted in the rate measured in units of s^{-1}. Upon experiment completion (300 s typical time) the normalised G-band area versus time dependence ('growth curve') was automatically fit with equation (2) (Supplementary Figure S2b), and the resulting kinetic parameters were written into the database along with experimental inputs. The last room-temperature spectrum (Supplementary Figure S2c) was acquired with 200 mW power and 30 s exposure:

$$G(t) = G_{\max}\left(1-\frac{1+\theta_0}{1+\exp((t-t_0)/\tau)}\right). \quad (2)$$

The updated database was transferred to the planner, and the cycle was repeated. This allowed up to 25 experiments in autonomous mode (without human intervention), updating the database after each experiment to build an improved model.

Electron microscopy characterisation was done in FEI Sirion SEM (FEI Company, Inc., Hillsboro, OR, USA) equipped with a field-emission gun and a through-lens detector (TDI) detector, at 4 keV accelerating voltage. Since pillars were on a silicon dioxide sublayer, charging could be severe. As a result, nanotubes present as sparse films (Figure 2a) often appeared in inverse contrast (dark lines on bright background). Thick nanotube mats (Figures 2b and c) tend to eliminate charging.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding from the Air Force Office of Scientific Research. The authors thank Dr Gordon Sargent of UES Inc. for help with scanning electron microscopy. The authors also thank Dr Mauricio Terrones and Dr Nestor Perea-Lopez of Pennsylvania State University for help with substrate preparation.

CONTRIBUTIONS

PN, DH and RR built the ARES system. PN, DH and FW ran the experiments. DH, FW and KD developed the ARES control software. MK, JP and RB developed and implemented the NMD-M3 experimental planner. PN, MK, JP and BM analysed the data and prepared the manuscript. BM supervised this work. All authors discussed the results and implications and commented on the manuscript.

COMPETING INTERESTS

The authors declare no conflict of interest.

REFERENCES

1. Graedel, T. E., Harper, E. M., Nassar, N. T. & Reck, B. K. On the materials basis of modern society. *Proc. Natl Acad. Sci. USA* **112**, 6295–6300 (2015).
2. American Chemical Society National Historic Chemical Landmarks. High performance carbon fibers. Available at http://www.acs.org/content/acs/en/education/whatischemistry/landmarks/carbonfibers.html.
3. Baker, M. Academic screening goes high-throughput. *Nat. Methods* **7**, 787–792 (2010).
4. Masimirembwa, C. M., Thompson, R. & Andersson, T. B. *In vitro* high throughput screening of compounds for favorable metabolic properties in drug discovery. *Comb. Chem. High Target Scr.* **4**, 245–263 (2001).
5. Giuliano, K. A., Haskins, J. R. & Taylor, D. L. Advances in high content screening for drug discovery. *Assay Drug Dev. Techn.* **1**, 565–577 (2003).
6. Shockcor, J. P. & Holmes, E. Metabonomic applications in toxicity screening and disease diagnosis. *Curr. Top. Med. Chem.* **2**, 35–51 (2002).
7. Dunlop, J., Bowlby, M., Peri, R., Vasilyev, D. & Arias, R. High-throughput electrophysiology: an emerging paradigm for ion-channel screening and physiology. *Nat. Rev. Drug Discov.* **7**, 358–368 (2008).
8. Potyrailo, R. *et al.* Combinatorial and high-throughput screening of materials libraries: review of state of the art. *ACS Comb. Sci.* **13**, 579–633 (2011).
9. Gebhardt, T., Music, D., Takahashi, T. & Schneider, J. M. Combinatorial thin film materials science: from alloy discovery and optimization to alloy design. *Thin Solid Films* **520**, 5491–5499 (2012).
10. Neuber, C. *et al.* Combinatorial preparation and characterization of thin-film multilayer electro-optical devices. *Rev. Sci. Instrum.* **78**, 072216 (2007).
11. Chan, E. M. *et al.* Reproducible, high-throughput synthesis of colloidal nanocrystals for optimization in multidimensional parameter space. *Nano Lett.* **10**, 1874–1885 (2010).
12. Cassell, A. M., Verma, S., Delzeit, L., Meyyappan, M. & Han, J. Combinatorial optimization of heterogeneous catalysts used in the growth of carbon nanotubes. *Langmuir* **17**, 260–264 (2001).
13. Noda, S., Tsuji, Y., Murakami, Y. & Maruyama, S. Combinatorial method to prepare metal nanoparticles that catalyze the growth of single-walled carbon nanotubes. *Appl. Phys. Lett.* **86**, 173106-1–173106-3 (2005).
14. Oliver, C. R. *et al. Robofurnace*: a semi-automated laboratory chemical vapor deposition system for high-throughput nanomaterial synthesis and process discovery. *Rev. Sci. Instrum.* **84**, 115105-1–115105-14 (2013).
15. Nikolaev, P., Hooper, D., Perea-Lopez, N., Terrones, M. & Maruyama, B. Discovery of wall-selective carbon nanotube growth conditions via automated experimentation. *ACS Nano* **8**, 10214–10222 (2014).
16. Spowart, J. E. Automated serial sectioning for 3-D analysis of microstructures. *Scr. Mater.* **55**, 5–10 (2006).
17. Jain, A. *et al.* Commentary: The Materials Project: a materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013).
18. Panchal, J. H., Kalidindi, S. R. & McDowell, D. L. Key computational modeling issues in Integrated Computational Materials Engineering. *Comput. Aided Design* **45**, 4–25 (2013).
19. Powell, W. B. & Ryzov, I. O. *Optimal Learning. Wiley Series in Probability and Statistics*, Vol. 841 (John Wiley & Sons, 2012).
20. Kelly, A. *et al.* Toward reliable off road autonomous vehicles operating in challenging environments. *Int. J. Robot. Res.* **25**, 5–6 (2006).
21. Schenker, P. S., Huntsberger, T. L., Pirjanian, P., Baumgartner, E. T. & Tunstel, E. Planetary rover developments supporting Mars exploration, sample return and future human-robotic colonization. *Auton. Robots* **14**, 103–126 (2003).
22. Dierks, T. & Jagannathan, S. Output feedback control of a quadrotor UAV using neural networks. *IEEE Trans. Neural Netw.* **21**, 50–66 (2010).
23. Yang, Y., Polycarpou, M. M. & Minai, A. A. Multi-UAV cooperative search using an opportunistic learning method. *J. Dyn. Syst-T ASME* **129**, 716–728 (2007).
24. King, R. D. *et al.* Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–252 (2004).
25. King, R. D. *et al.* The automation of science. *Science* **324**, 85–89 (2009).
26. Sparkes, A. *et al.* Towards robot scientists for autonomous scientific discovery. *Autom. Exp.* **2**, 1-1–1-11 (2010).
27. Williams, K. *et al.* Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases. *R. Soc. Interface* **12**, 20141289 (2015).
28. Bilsland, E. *et al.* Yeast-based automated high-throughput screens to identify anti-parasitic lead compounds. *Open Biol.* **3**, 120158-1–120158-13 (2013).
29. De Volder, M. F. L., Tawfick, S. H., Baughman, R. H. & Hart, A. J. Carbon nanotubes: present and future commercial applications. *Science* **339**, 535–539 (2013).
30. Harvey, S. E. Carbon as conductor: a pragmatic view. Proceedings of the 61st IWCS Conference, http://www.iwcs.org/archives/56333-iwcs-2012b-1.1584632/t-001-1.1585113/f-013-1.1585221/12-4-1.1585228/12-4-1.1585229 (2012).

31. Rao, R., Liptak, D., Cherukuri, T., Yakobson, B. I. & Maruyama, B. *In situ* evidence for chirality-dependent growth rates of individual carbon nanotubes. *Nat. Mater.* **11**, 213–216 (2012).
32. Rao, R. *et al.* Revealing the impact of catalyst phase transition on carbon nanotube growth by *in situ* Raman spectroscopy. *ACS Nano* **7**, 1100–1107 (2013).
33. Rao, R., Islam, A. E., Pierce, N., Nikolaev, P. & Maruyama, B. Chiral angle-dependent defect evolution in CVD-grown single-walled carbon nanotubes. *Carbon* **95**, 287–291 (2015).
34. Islam, A. *et al.* Photo-thermal oxidation of single layer graphene. *RSC Adv.* **6**, 42545–42553 (2016).
35. Liaw, A. & Weiner, M. Classification and regression by Random Forest. *R. News* **2**, 18–22 (2002).
36. Scrucca, L. GA: a package for genetic algorithms in *R. J. Stat. Softw.* **53**, 1–37 (2012).
37. Carvalho, D. R. & Freitas, A. A. A hybrid decision tree/genetic algorithm method for data mining. *Inf. Sci.* **163**, 13–35 (2004).
38. Krein, M., Huang, T. W., Morkowchuk, L., Agrafiotis, D. K. & Breneman, C. M. in *Statistical Modelling of Molecular Descriptors in QSAR/QSPR* (eds Dehmer, M., Varmuza, K., Bonchev, D. & Emmert-Streib, F.) Ch. 2 (Wiley-Blackwell, 2012).
39. Weaver, S. & Gleeson, M. P. The importance of the domain of applicability in QSAR modeling. *J. Mol. Graph. Model.* **26**, 1315–1326 (2008).
40. Futaba, D. *et al.* Kinetics of water-assisted single-walled carbon nanotube synthesis revealed by a time-evolution analysis. *Phys. Rev. Lett.* **95**, 056104-1–056104-4 (2005).
41. Chen, G. *et al.* A sweet spot for highly efficient growth of vertically aligned single-walled carbon nanotube forests enabling their unique structures and properties. *Nanoscale* **8**, 162–171 (2016).
42. Wood, R. F., Pannala, S., Wells, J. C., Puretzky, A. A. & Geohegan, D. B. Simple model of the interrelation between single- and multiwall carbon nanotube growth rates for the CVD process. *Phys. Rev. B* **75**, 235446-1–235446-8 (2007).
43. Einarssona, E., Yoichi Murakamia, Y., Masayuki Kadowakia, M. & Shigeo Maruyama, S. Growth dynamics of vertically aligned single-walled carbon nanotubes from in situ measurements. *Carbon* **46**, 923–930 (2008).
44. Vinten, P., Lefebvre, J. & Finnie, P. Kinetic critical temperature and optimized chemical vapor deposition growth of carbon nanotubes. *Chem. Phys. Lett.* **469**, 293–297 (2009).
45. Balkanski, M., Wallis, R. F. & Haro, E. Anharmonic effects in light scattering due to optical phonons in silicon. *Phys. Rev. B* **28**, 1928–1934 (1983).

Supplementary Information accompanies the paper on the *npj Computational Materials* website (http://www.nature.com/npjcompumats)