Accelerating deep learning discovery of new thermoset shape memory polymers

Cheng Yan, Xiaming Feng, Louisiana State University; Collin Wick, Andrew Peters, Louisiana Tech University; Guoqiang Li, Southern University

Award Title:	RII Track-1: Louisiana Materials Design Alliance (LAMDA)
NSF Award Number:	NSF OIA-1946231
Principal Investigator:	Michael Khonsari
Lead Institution Name:	Louisiana State University
Award Start Date:	07/01/2020
Award End Date:	06/30/2025
Highlight Submission Date:	3/11/2021

What is the outcome or accomplishment?

Researchers with the Louisiana Materials Design Alliance (LAMDA) established a machine learning framework to predict the recovery stress of thermoset shape memory polymers (TSMPs) and to discover new TSMPs with superior recovery stress. The team leveraged a new linear notation computer language for the digital representation of polymers, called BigSMILES, to fingerprint complex TSMP structures and establish structure-property correlations using a small training dataset. This information helped to identify two new TSMPs predicted to have high recovery stress, which was synthesized in the lab to validate the model predictions. Finally, they explored a chemical space with 4,459 possible TSMPs and screened 14 mostly unknown TSMPs with higher recovery stress than any TSMPs in the training dataset. One of 14 TSMPs was modeled by molecular dynamics simulation and found to have calculated recovery stress in agreement with predicted values.

What is the impact?

To our knowledge, this study is the first to discover TSMPs with high recovery stress by leveraging machine learning. Using a small dataset of about 100 molecules, this method has the potential to greatly increase the ability to explore chemical space and bring remarkable advancements over previous materials discovery methods.

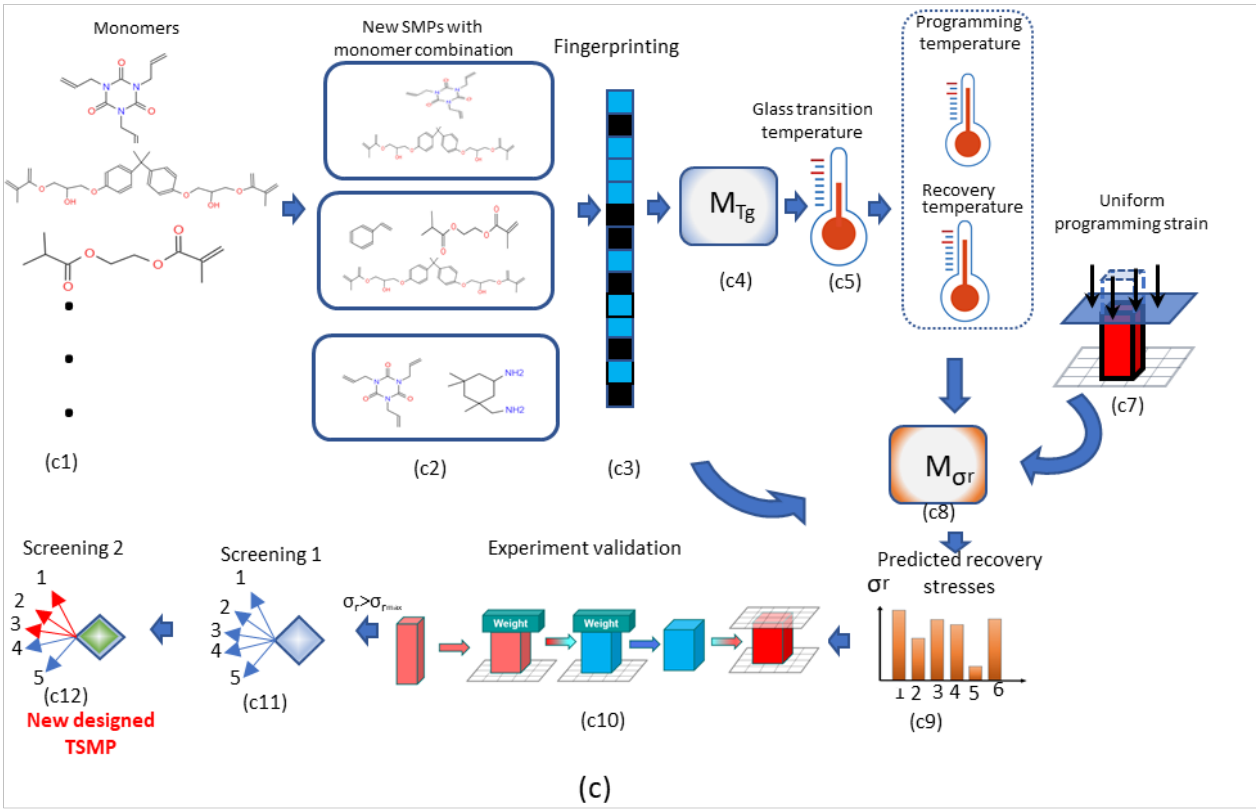
a) A bottleneck for current TSMPs persists in their low recovery stress in their rubbery state, limiting their applications as actuators or as crack closing devices in self-healing applications.
b) Due to the time it takes to synthesize new TSMPs, the traditional trial-and-error method for materials discovery needs a long period of time, deep domain knowledge and skills, and some luck. This new method overcame these limitations and quickly discovered 14 new TSMPs.
c) The machine learning approach is at least hundreds of times faster than the traditional molecular or atomistic computational approaches, such as molecular dynamics simulation and density functional theory based electronic calculation.
d) Traditionally, predictions of thermomechanical behaviors of TSMPs, such as recovery stress, rely on multi-parameter constitutive models, and most of the parameters need to be

determined by curve-fitting. This machine learning model, on the other hand, only depends on the basic chemical structures and so can be applied nearly universally.

What explanation/background does the lay reader need to understand the significance of this outcome?

TSMPs are a new class of smart polymers, which, after deformation, can maintain their deformation nearly permanently until a stimulus, such as heat or an electric field, is applied. This stimulus causes the deformed TSMP to restore its original shape. TSMPs have found many applications, such as crack self-healing in lightweight structures, stents in medical applications, and artificial muscles in soft robots. However, a persisting critical limitation of existing TSMPs is their very low recovery stress in their rubbery state, usually less than a few Newtons per square millimeter. This recovery stress is too low for some critical applications, such as the crack closing. It is also too low to compete with shape memory alloys, which may have tens to hundreds of Newtons per square millimeters in recovery stress. Unfortunately, the chemical space constituting TSMPs is very large, and using trial and error approaches is not sufficient to identify new TSMPs with higher recovery stress. Therefore, machine learning is a natural choice. As compared to conventional polymers—which have already had a large database in Materials Genome Initiative—only a limited number of TSMPs are available for training and fingerprinting. In this study, we successfully overcame this limitation and discovered 14 new TSMPs. This machine learning approach can be expanded to discover other types of materials.

Mr. Cheng Yan, a Ph.D. student in the Department of Mechanical & Industrial Engineering at Louisiana State University, is designing and encoding the machine learning pipeline.



Pipeline for new TSMPs discovery: first, collecting monomer set in the two datasets (c1), and then automatically generating random combinations of monomers and crosslinkers, which produce new TSMPs (c2). Next, fingerprinting is performed for these new TSMPs (c3) and input into the glass transition temperature model (c4) to predict the corresponding glass transition temperature (c5). Then, programming temperature T_{tr}, recovery temperature (c6) (approximated from glass transition temperature, i.e., $T_{tr} = T_g + 20$ °C), presumptive uniform strain (c7), and fingerprints of the newly formed TSMP structures (c3) are input into recovery stress model (c8) to predict the corresponding recovery stress σ_s (c9), which is validated by experiments (c10). Finally, two-step screening processes were conducted. First, if the predicted recovery stress is greater than the maximum recovery stress in the training data, then the corresponding TSMPs are recorded (c11); second, by employing the prior knowledge, promising TSMPs with higher recovery stress than that in the training dataset are be further screened (c12).