# Semi-Automated Discovery in Zariski Spaces (A Proposal)

**Simon Colton, Roy McCasland, Alan Bundy and Toby Walsh,**
**Institute of Representation and Reasoning**
**Division of Informatics, University of Edinburgh**
**80 South Bridge, Edinburgh. EH1 1HN. UK.**

We discuss the proposed application of the HR discovery system (and the various pieces of mathematical software it uses) to discovery tasks in the domain of Zariski Spaces. We propose an incremental approach using HR in a semi-automated, cross-domain fashion in domains of increasing difficulty related to Zariski spaces. This work is to be funded by an EPSRC 1-year visiting fellowship to employ Roy McCasland.

## Zariski Spaces

Zariski Spaces were introduced in 1998 [MMS98]. In order to understand these spaces, one needs to first understand Zariski Topologies. To this end, let $R$ be a commutative ring with unity, and let $specR$ denote the collection of prime ideals of $R$. Now for each (possibly empty) subset $A$ of $R$, let the *variety* of $A$ be given by: $V(A) = \{P \in specR : A \subseteq P\}$. It is easily shown that the collection of all such varieties constitutes (the closed sets of) a topology, called the Zariski Topology on $R$, which we denote by $\zeta(R)$. It turns out that every topology is a semiring, if one takes the operations of addition and multiplication to be set-theoretic intersection and union respectively. Now let $M$ be an $R$-module, and repeat the above process. That is to say, let $specM$ denote the collection of all prime submodules of $M$, and for each subset $B$ of $M$, let the variety of $B$ be given by: $V(B) = \{P \in specM : B \subseteq P\}$. Finally, let $\zeta(M)$ represent the collection of all varieties of subsets of $M$. Then one can show that while $\zeta(M)$ seldom forms a topology, it does form a semimodule over the semiring $\zeta(R)$, where the operation in $\zeta(M)$ is taken to be intersection, and scalar multiplication is given by $V(A)V(B) = V(RAB)$.

There are reasons to suppose that Zariski Spaces might turn out to be of considerable importance in mathematics. For instance, Zariski Topologies and the study of varieties have played an enormous role in the development of Algebraic Geometry, and in particular, the Hilbert Nullstellensatz, which is one of the fundamental results in Algebraic Geometry. It is certainly possible that Zariski Spaces could have a similar impact on some branch of mathematics. Furthermore, some preliminary results suggest a possible connection between a certain concept in Zariski Spaces, called subtractive bases, and direct sum decompositions within a large class of modules. The search for direct sum decompositions has been a major undertaking in mathematics for some time, and has so far proven quite intractable, except in special cases. The study of semimodules in general has already yielded many applications to computer science [Gol92], and since Zariski Spaces are first and foremost semimodules, it is possible that their study will promote further advances in theoretical computer science.

## Proposed Discovery Methods

The HR program [Col00] has been successful in making discoveries in number theory [Col99] and algebraic domains [CM01]. HR is comprised of four modules which generate four types of information, namely objects of interest, concepts which classify those objects, conjectures which relate the concepts and proofs which explain the conjectures. HR calls third party software to achieve various tasks, including computer algebra systems, constraint solvers and model generators to generate objects of interest (for both exploration and counterexamples) and theorem provers to prove conjectures. At present, HR is autonomous, i.e., the user sets some parameters for the search it will perform, then HR builds a theory and the user employs various tools to extract information about the theory. We propose to extend the theory of theory formation upon which HR is based by enabling any of the four modules to be replaced on occasion by human intervention. That is, we intend to make HR semi-automated by allowing the user to provide proofs and counterexamples to conjectures HR makes and to specify related concepts and conjectures to base the theory formation around. Alongside the development of HR's functionality, we also intend to develop the application to Zariski spaces. Zariski spaces represent a higher level of complexity than the domains in which HR has so far been applied, and the new application will require an incremental approach whereby (i) HR is enabled to form theories about increasingly complicated domains related to Zariski spaces (ii) testing is performed to see if HR invents various concepts and conjectures required for it to proceed and (iii) theory formation is centred around the important concepts and the results analysed for any discoveries. The proposed route to Zariski spaces is via: semigroups, semirings and semimodules, followed by groups, rings and modules and finally, using a cross domain approach, Zariski spaces.

It is unclear at the moment how HR will interact with the user and with third party pieces of mathematical software on this project. It seems likely that HR will be used for an initial exploration of each domain, to: (a) invent core concepts (b) prove some fundamental theorems using a theorem prover (HR has access to Otter, E, Spass and Bliksem through the MathWeb Software Bus), and (c) generate some examples of the concepts using a model generator or computer algebra system (HR uses MACE and Maple, also possible through MathWeb). Following the initial investigation, the user will both prune uninteresting concepts and specify which concepts should be emphasised in the next theory formation session. The user will be more involved in that session, choosing to prove theorems, provide counterexamples and direct the search where appropriate. By improving HR to enable such an interaction, we hope this approach will lead to the discovery of new results about Zariski spaces.

## References

[CM01] S Colton and I Miguel. Constraint generation via automated theory formation. In *Proceedings of CP-2001*, pages 575–579, 2001.

[Col99] S Colton. Refactorable numbers - a machine invention. *Journal of Integer Sequences*, 2, 1999.

[Col00] S Colton. *Automated Theory Formation in Pure Mathematics*. PhD thesis, Department of Artificial Intelligence, University of Edinburgh, 2000.

[Gol92] J Golan. *The theory of Semirings with Applications in Mathematics and Theoretical Computer Science*. John Wiley and Sons, 1992.

[MMS98] R McCasland, M Moore, and P Smith. An introduction to Zariski spaces over Zariski topologies. *Rocky Mountain Journal of Mathematics*, 28:1357–1369, 1998.