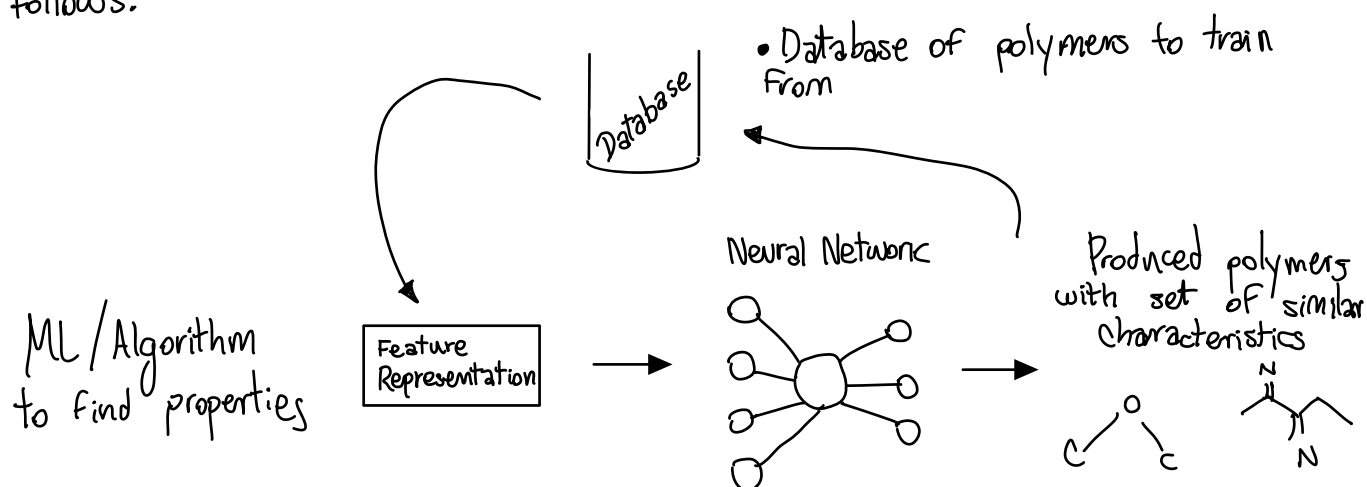


Architecture Proposal: So for this project and this plan for architecture proposal, I will try to draw a diagram over the small couple of things that compose the entire project, and then from there, I will determine my course of action during the weekend.

Goal: Build a ML architecture that will allow for the training of polymers for certain properties. The pipeline would look as follows:



Representation
UI/Interface

→ 3D representation of molecules

Things to find:

Database to use/
how to get data
to train from?

Problem: Databases are expensive and the ones available might not have the properties we need?

N/A

↓
Data
Augmentation;
JUST LOOK HARDER BRO

It will inevitably
require the use of
AWS or MIT
Supercloud

What ML Alg to use?
Deep RL? MO? Trees?

→ Most likely something that has
to do with Deep RL

Which Feature Representation is the most Feasible?

Using 3D-structures as described in the GeoMol or Deep RL paper.

Problem: how do we know the properties of polymers from the properties of monomers?

Maybe start with smaller steps; focus on a single kind of molecule (e.g. cellulose acetate).

What's the problem then?

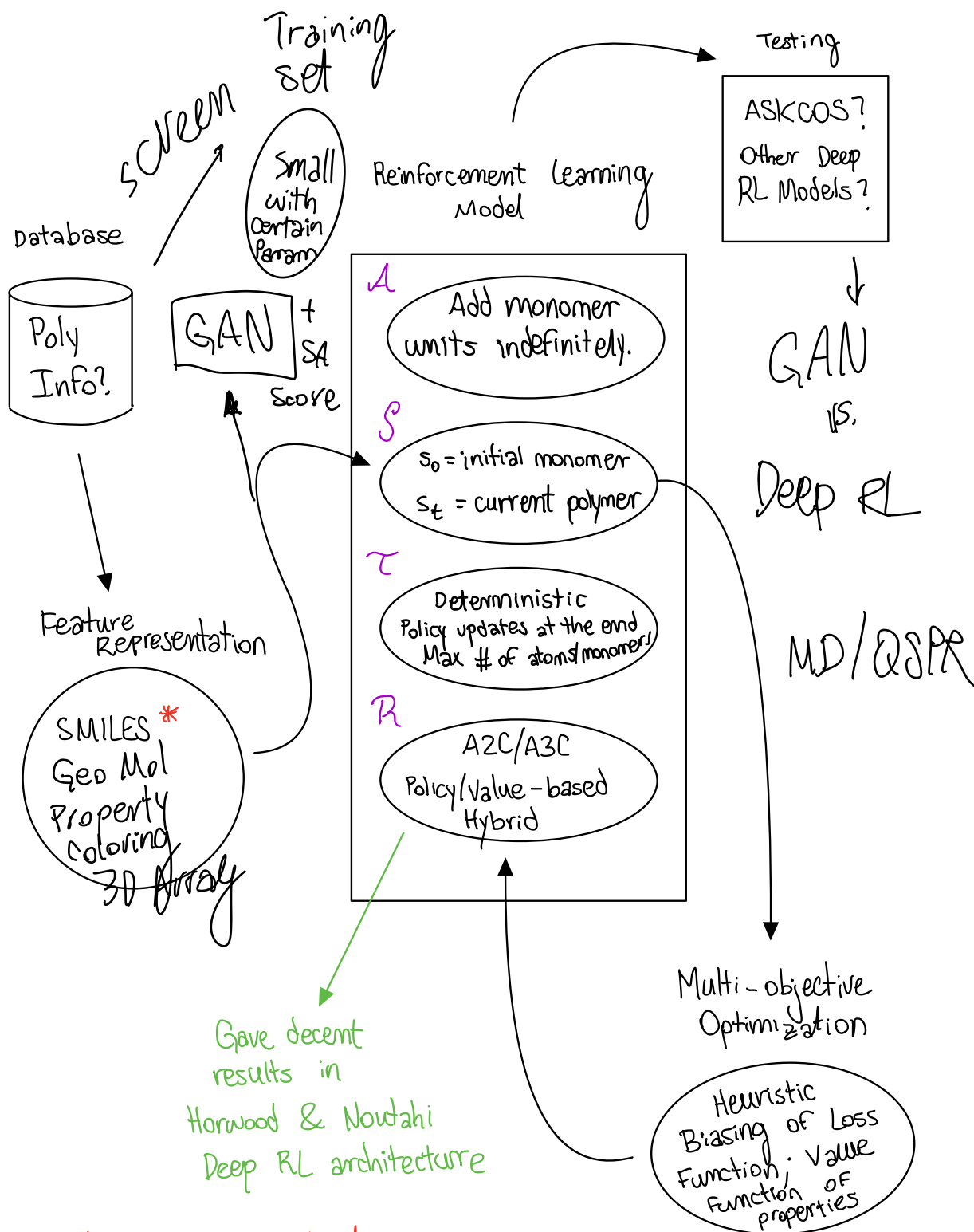
⇒ It's hard to give a proper guess as to what's the step forward without ample knowledge of the literature beforehand. I must read more about the literature and focus on:

- Performance
- Computational Expense
- How much data is available
- How feasible it is (make sure it is not an absolutely crazy idea and that it can work under current circumstances)

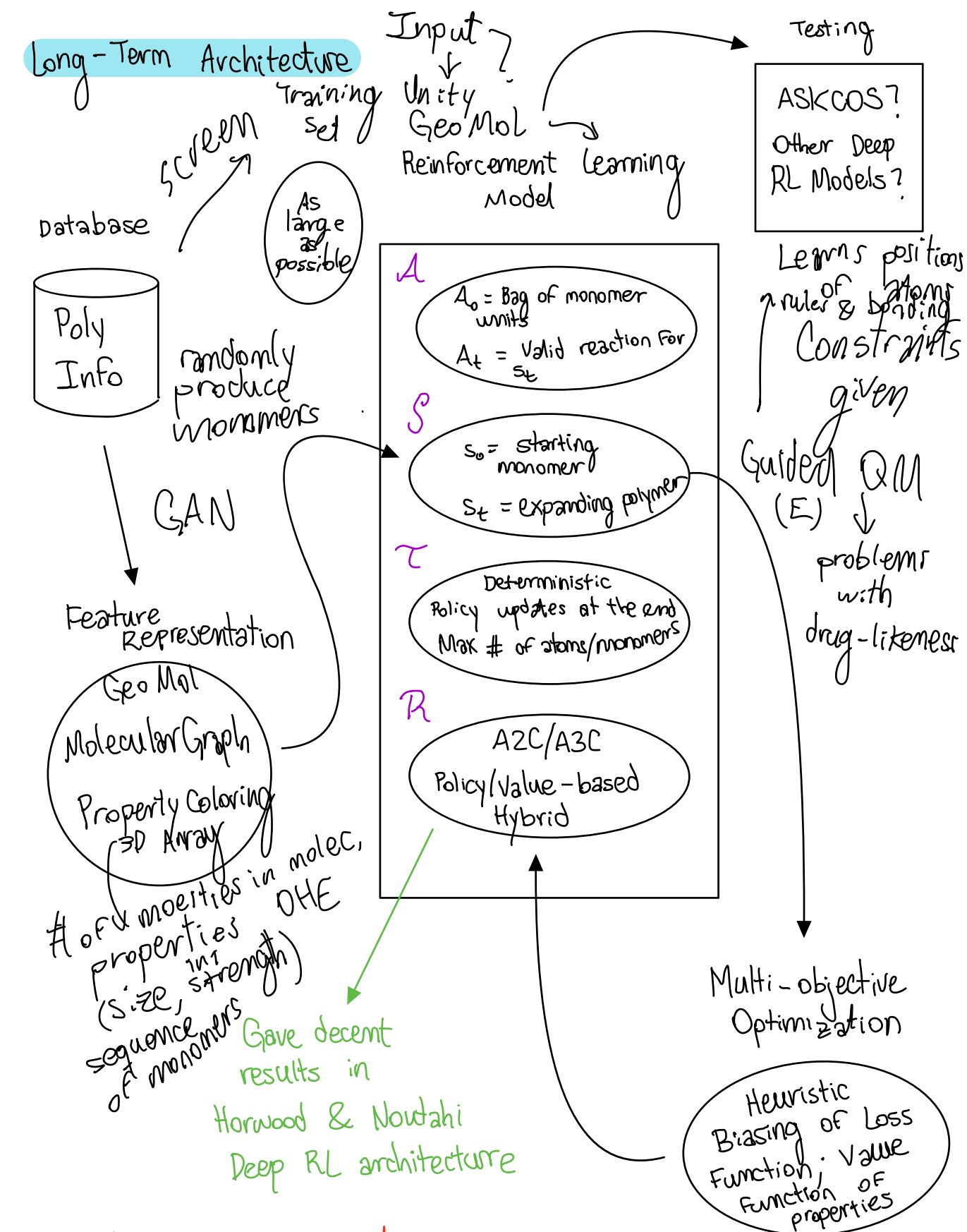
For Example:

- Change representation from 3D structure to Cartesian
- Input only bond angles (???)
- Etc

Short-Term Architecture aka. Low Data



* = Might be inefficient/useless



* = Might be inefficient / useless

*Data Load Model \Rightarrow Poly Info 1mil (not really
to be trusted

Poly Info (waiting to
obtain it)

↓
It's on Github, it could
be useful for a toy
Data Set.

Specific vs. General One

↓

Objective: Produce molecules that resemble this dataset of well-known molecules
(the process works)

↘

Objective: Generative model for polymers; optimizing for multiple objectives.