Using Motion Planning to study Self-Assembly by Shape Recognition[*]

Urmi Majumder

Abstract

In this project we study self-assembly via shape recognition in a two-tile system using a motion planning technique called *probabilistic roadmap*. The latter is a very useful method to check the connectivity of a collision free path in a high dimensional space. But it can also be applied to capture pertinent features of a continuous function over high dimensional space. Thus in this project we use probabilistic roadmaps to study the potential energy landscape of a two-tile self-assembly. Our probabilistic roadmap tries to capture the connectivity of the low energy subset of the tile configuration space in the form of a network of weighted pathways. In this work we describe general techniques to compute and query our probabilistic roadmap.

1 Introduction

Self assembly is the autonomous organization of components into structures without human intervention. Self assembly is a phenomenon that is prevalent on all scales, from molecules to galaxies. Examples of self-assembling system include weather patterns, solar systems, self-assembled monolayers etc. The most well-studied subfield of self-assembly is molecular self-assembly. However, most of the interesting applications of self-assembling processes can be found at a larger scale. Also we have a lot of experimental convenience and control for meso-scale self-assembly. In fact, massively parallel self-assem bling systems present a promising alternative to conventional manufacturing (pick and place assembly). Some typical examples of mesoscale assembly has been given in figure 1.

Thus in this project we've looked into self-assembly of macroscopic $2D$ and $3D$ "tiles" by shape recognition. The main motivation for studying mesoscale $3D$ self-assembly[1] is the fabrication of functionalized $3D$ electronic systems. There has been several demonstrations of meso-scale self-assembly using capillary forces[2], electrostatic[3] and magnetic attraction[4],

---

[*] This is CPS234 term project

shape recognition[5] and several others. However, most of the demonstrations are for simple periodic lattices. In order to build functional units we need to incorporate some complexity into our designs. This can be done by using some encoding techniques or giving our assembly units also known as tiles complementary shapes. One of the main problem with binary encoding techniques is the extent of partial matches among tiles. Self-assembly by shape recognition scores better in this regard.

In this project we present a framework for studying the pathways for self-assembly by shape recognition and potential energy landscapes which is based on techniques widely used in robot motion planning community. The goal of motion planning is to compute a sequence of valid intermediate states to transform a given initial state into some desired final state. Although self-assembly by shape recognition appears to be totally unrelated to the traditional motion planning problem, the algorithms there are often described using an abstraction called configuration space which is sufficiently general to apply to our problem. Briefly, the configuration space, or C-space of a movable object is the space consisting of all positions and orientations of the object.

In this project, we concentrate on the application of a motion planning technique called Probabilistic Roadmaps(PRM)[6] to study self-assembly by shape recognition. The primary motivation for studying the potential energy landscapes is to have better tile designs and hence ensure a more correct assembly. We have selected PRM because of its proven success in exploring high-dimensional configuration spaces. In fact, PRM has been used to study the problem of protein-ligand binding[7] which is somewhat similar to the work we present here.

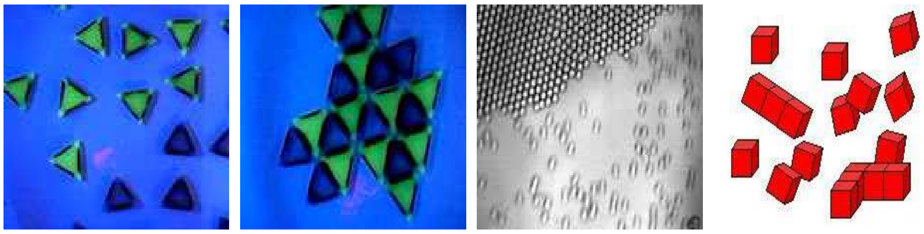

Figure 1: **Examples of Mesoscale Assembly**(*from left to right*)(a)+(b) Triangles, made of polydimethylsiloxane with magnetized strontium ferrite along the edges, self assemble on the surface of water, (c)SA of 5 μm glass nanospheres on a mica surface, (d) Self-assembling meso blocks in a liquid environment

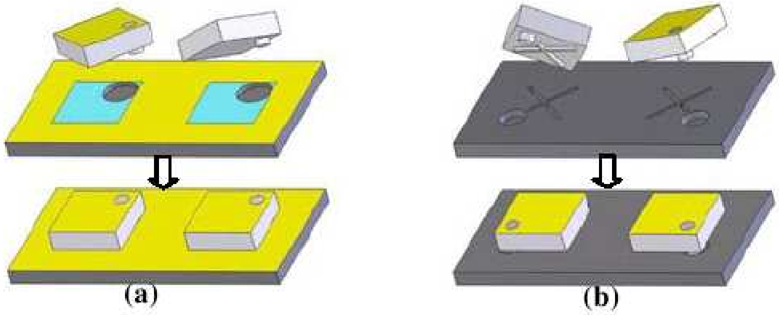

Figure 2: Schematic diagram of feature directed self-assembly (a)Each part has a single circular peg uniquely aligned with its receptor site having a circular trench (b)Each part has a circular peg and a cross peg uniquely aligned with its receptor site having a circular trench and a cross trench(adopted from [5])

2 Previous Work

The first demonstration of self-assembly dates back to early 1990s when Leonard Adelman at USC first used the tools of molecular biology to solve an instance of the directed Hamiltonian Path problem. A small graph was encoded in molecules of DNA and the "operations" of computation were performed using standard protocols and enzyme. The experiment demonstrated the feasibility of carrying out computation at the molecular level. [8]. There has been several demonstrations since then using a variety of "tile types" and driving forces. The most significant demonstration of meso-scale self-assembly by shape recognition was by Karl Böhringer's group[5] where they demonstrated two different feature directed and uniquely orienting self-assembly process.: one is a semi-dry process based on gravity driven self-alignment and another one is a completely dry process driven by two stage shape recognition[Figure 2]

As we mentioned in the previous section, we 'll borrow techniques from the field of robot motion planning to study macroscopic self-assembly by shape recognition. The basic problem in robot motion planning is to generate a collision-free path for a rigid or articulated object ie the robot among static obstacles[Figure 3]. The input to the problem are

- Geometry of the obstacles and the robot
- Kinematics of the robot
- Initial and goal configurations

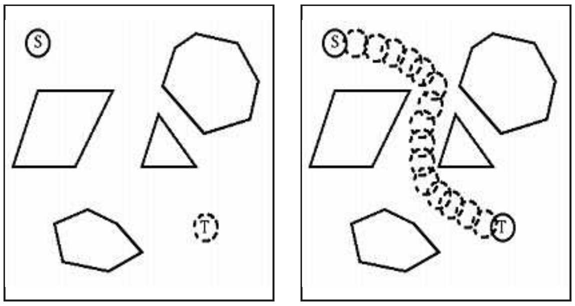

Figure 3: Robot Motion Planning-input(left), output(right)

and the output of the problem is a contiguous sequence of collision-free configurations connecting the initial and goal configurations.

Path planning for robots have been studied extensively over the past three decades.[9]. However, the complexity of complete path planning methods in high dimensional configuration spaces has led to the development of heuristic methods that embed weaker notions of completeness , for instance probabilistic completeness. One unique approach among all those heuristic methods is that of applying a roadmap approach which is essentially a network of paths in free space. There has been several roadmap techniques viz. visibility graph[10], Voronoi diagram[11], and silhouette[12] methods. However, they compute in a single shot a roadmap that completely represents the connectivity of the free C space. which makes them useful for computing roadmap only in low dimensions. *Probabilistic roadmaps*[6] on the other hand, emphasizes on efficiency and thus can be effectively applied to high dimensional C space because it builds the roadmap incrementally. The method proceeds in two phases: a learning phase and a query phase. In the learning phase, a probabilistic roadmap is constructed and stored as a graph whose nodes correspond to collision-free configurations and whose edges correspond to feasible paths between these configurations. These paths are computed using a simple and fast local planner. In the query phase, for any given start and goal configurations of the robot, the roadmap is searched for a path joining these two nodes.[Figure 4, 5].

In the recent years probabilistic roadmaps have been used to capture the pertinent features of continuous functions over high dimensional spaces., for instance computing the energetically favorable motions of biomolecules.[7, 13] Here a molecule is modeled as an articulated object moving in an energy field. The set of all its $3D$ placements and orientation defines its configuration space C. A probabilistic roadmap tries to capture the connectivity of the low energy subset of the configuration space in the form of a network of

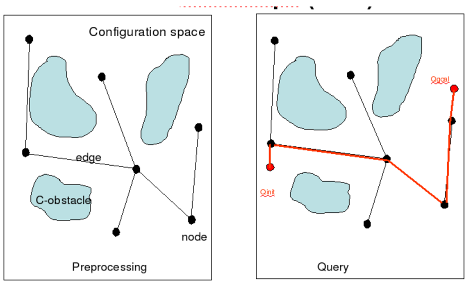

Figure 4: Probabilistic Roadmaps for path planning in high dimensional space

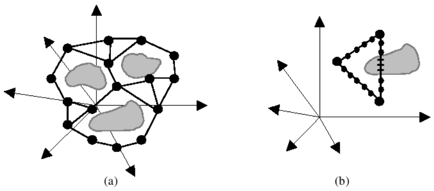

Figure 5: (a)Probabilistic Roadmaps for a six dimensional space, (b)Local Path Planner

weighted local pathways[Figure 6]. The weight of a pathway measures the energetic difficulty for the molecule to move along it.

Thus the probabilistic roadmaps which are an effective tool for verifying the connectivity of a collision-free subset of a high dimensional robot configuration space, easily extends to capture the main features of a continuous function over the high dimensional space. Hence this general paradigm can be used to study the *energy landscape* over the configuration space of the two-tile self-assembling system. Our hope is that our study will help us understand the correlation between tile shapes and ease of assembly better.

3 Probabilistic RoadMaps for Tile Assembly by Shape Recognition

3.1 Goals

The primary goals of our study are

- Examine the dynamics of one of the tile as it approaches the other tile with complementary shape in order to get bound to it.

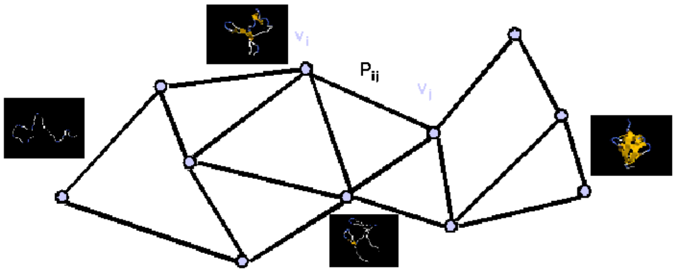

Figure 6: Probabilistic Roadmaps for the protein folding problem(Adopted from [7])

- Check whether an energetically favorable path exists between any randomly generated start configuration for the tile and its final bound configuration
- Study the energy barrier for tile-assembly by shape recognition
- Study how the energy barrier varies with different complementary shapes.
- Observe change in the potential energy landscape with the variation in the depth of the binding site
- Examine the change in depth of the energy barrier with change in the width of binding site.
- Distinguish a true binding site from a false binding site(due to partial mismatches or physical obstacle)

3.2 Benefits

A very common method for studying thermodynamic and kinetic properties of biomolecule docking(for instance protein-ligand binding) is to use computationally intensive simulation techniques like Molecular dynamics[14] and Monte Carlo Simulation[15]. While the former attempts to simulate the true dynamics of the system using Netwon's equations of motion, the latter uses a randomized approach for generating successive ligand configurations and compute the thermodynamic properties by averaging over all samples that were generated[Figure 7]. Both these technniqes are applicable for studying energy landscape in a tile-assembly model too. However, an approach based on motion planning called *probabilistic roadmaps*[9, 6] will significantly improve the speed and efficiency of the simulation. The idea is to alleviate

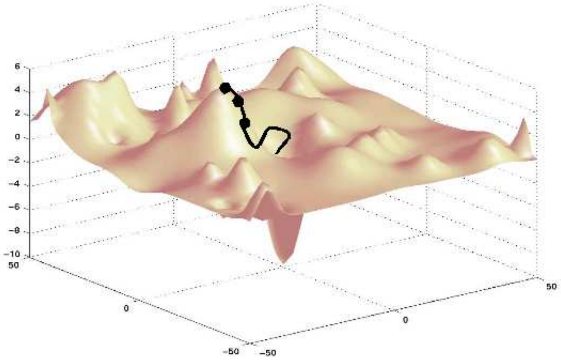

Figure 7: Monte Carlo Simulation find only one path at a time on this energy landscape.

the time dependency of molecular dynamics method and the Markov dependency of the Monte Carlo methods by sampling from the space of all possible paths that a tile may take as it binds to the other tile. Precisely we guess several possible intermediate configurations of the tile B and obtain a distribution of energetically favorable paths to the binding site via the intermediate configurations. In addition, we measure and compare the difficulty in traversing each of these path and hence obtain an estimate for the energy barrier. We can also obtain estimates for the rates of binding and dissociation of the two tiles.

3.3 Setup

In this system we study the assembly of two tiles moving in a rectangular air bed. The two tiles have complementary shapes with tiny magnetic dipoles arranged on the tile in such a manner that when the two tiles assemble via complementary shapes, the binding force of such an assembly is maximum[Figure 8]. For simplicity, we'll study the two dimensional case first and then analyze the three dimensional two-tile system. However, a three dimensional multi-tile system which is of main interest to us. An example assembly by magnetic dipole interaction only and not using any shape recognition has been shown in figure 9

3.4 Configuration Space

Since we are mostly interested in the energy barrier in tile-assembly, in the two tile system, we 'll fix one tile(say A) and allow the other tile(say B) to move. In two dimensions(*2D*), each tile will have three degrees of freedom(dof) -two translational(along x and y axes) and one rotational(where θ is the angle of rotation). For A, since its position and orientation is fixed

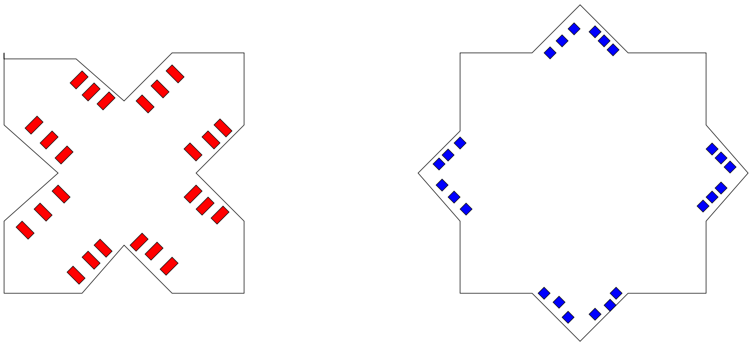

Figure 8: An Example of tiles with complementary shapes

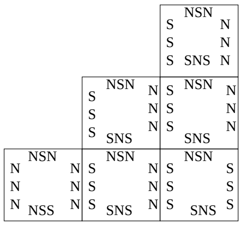

Figure 9: An example of Magnetic Assembly

we'll assign a particular value for each dof. We do this for the center of mass of the tile and as a result the position and orientation of any other point on the tile gets fixed. We define a configuration by assigning a value to each dof. The set of all configurations of the tile defines the tiles's configuration space C. As mentioned earlier, the tiles move in a rectangular air bed whose dimensions($a \times b$ sq units), so the bounds on each of the dimensions of the configuration space is $x \in [0, a]$, $y \in [0, b]$, $\theta \in [0°, 360°]$. We keep both a and b relatively small because we are mostly interested in the energy landscape around the binding site. We randomly choose a value for x, y, θ for tile A and time-dependent position and orientation of tile B will be discussed later. Note that the initial configuration of tile B is chosen at random, similar to the procedure for tile A. The final configuration of tile B corresponds to its position and orietation when it is matched to tile A via its complementary shape.

Thus $C \subseteq \mathbb{R}^3$ describes the conformational space of tile B[Figure 10]. Had the tiles been three dimensional then we would have six degrees of freedom and hence would have dealt with a six dimensional conformational space C[Figure 10]. In fact, if we remove the constraint that tile A has to be fixed, then for a three dimensional two tile system it would have been twelve

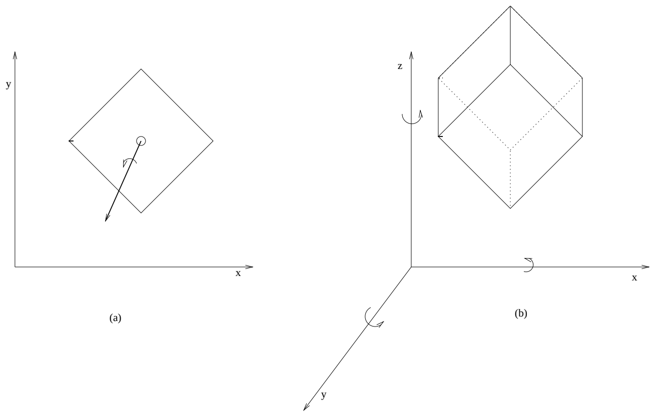

Figure 10: Degrees of Freedom of the tile in two dimensions(eft) and in three dimensions(right)

dimensional conformational space accounting for a total of tweleve degrees of freedom. In the most general scenario with n tiles our conformational space is in $\mathbb{R}^{6n}$.

3.5 Energy Computation

Let $E : C \longrightarrow \mathbb{R}$ be the potential energy field over C. This is essentially the potential energy of tile B due to the magnetic dipole interaction[Figure 11]. Computing the magnetic potential for every pair of dipole interaction is very expensive. Also appropriate shielding[Figure 11] of individual dipole reduces interdipole interaction to a large extent. The field is maximum at the "open" end and simulation shows that the flux lines tend to pass through the nearest available bar magnet of opposite polarity. Thus magnetic shielding reduces the intra-tile dipole interaction to a large extent. For instance in Figure 11, the shielded dipoles on the sides of the tiles, facing each other has the maximum field strength at the open ends while the remaining unshielded dipoles have a much spread out effect. In our system, each dipole is appropriately shielded except at the open end.. This will ensure minimum "crosstalk".

For simplicity we'll assume until the distance between tile B and tile A is below the threshold distance τ, the potential energy of tile B is only due to its own dipoles. Below τ, only the dipoles on the face of A nearest to B affect the potential energy of B. This simplification is valid because the ratio of dipole size to tile size is considerably small and the shielding effectively isolates the bar magnets on each side. It has to be borne in mind however, that though for macroscopic assembly one needs external

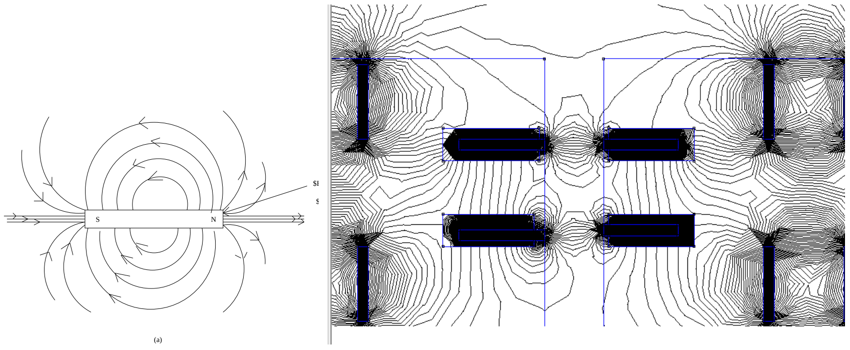

Figure 11: Flux lines of a bar magnet(left), effect of magnetic shielding(right)

driving force(for instance fan force for this particular experiment), but in this particular analysis we are interested in the behavior of the tile just before binding takes place. Under such a circumstance short-ranged force(in this case magnetic force) is more significant than the long-ranged force(in this case fan force). Hence while computing the potential energy we're assuming without much error that the sole contribution to potential energy is from the magnetic scalar potential.

The potential energy E of tile B at conformation $c \in C$ is

$$E(c) = E_A(c) + E_B$$

where $E_A(c)$ is the energy of tile B due to the dipoles on the side of A facing B when B is at c and $E_A(c)$ is mathematically given by,

$$E_A(c) = \sum_{i=1}^{n} m.B_i(c)$$

m is the dipole moment of the bar magnet, $B_i(c)$ is the magnetic field of the i^{th} dipole of A at the center of mass C_B of tile B. Formally,

$$B_i(c) = \frac{\mu_0}{4\pi}\frac{3\hat{r}(\overrightarrow{m}.\hat{r_i} - \overrightarrow{m})}{\overrightarrow{r_i}^3(c)}$$

where $\overrightarrow{r_i}(c)$ is the distance between C_B and center of mass C_{D_i} of the i^{th} dipole D_i on tile A. E_B is the energy of tile B due to its own dipoles and is defined in a similar manner

$$E_B = \sum_{i=1}^{n} m.B_i$$

B_i is the magnetic field of the i^{th} dipole of B at C_B where

$$B_i = \frac{\mu_0}{4\pi} \frac{3\hat{r}(\overrightarrow{m}.\hat{r_i} - \overrightarrow{m})}{\overrightarrow{r_i}^3}$$

$\overrightarrow{r_i}$ is the distance between C_B and center of mass C_{D_i} of the i^{th} dipole D_i on tile B.

3.5.1 Milestone Generation

In our PRM for tile-assembly, we assign each dof of tile B to a separate dimension in C space. Hence any configuration of tile B can be obtained by assigning random values to each coordinate in this high dimensional space. The milestones or the vertices of our PRM can be constructed by picking up a value uniformly at random for each conformational parameter viz. x, y, θ in two dimensions within its given range of possible values. Recall that, $x \in [0, a]$, $y \in [0, b]$, $\theta \in [0°, 360°]$.

Uniform sampling distribution is the simplest, but as we move to three dimensional tile assembly or relax the constraints on the tiles, the dimensionality of the conformational space increases and uniform sampling has to be very dense to reliably characterize the important features of the potential energy landscape. However, we can improve upon the current sampling strategy because we know the bound configuration of the tile. Thus, given the conformation of tile A, the final conformation of tile B can be obtained. So we can modify our sampling strategy so that its biased around the final conformation of tile B For instance we can have a set of normal distributions around the final conformation of tile B with a range of standard deviations from small to large. The small standard deviations will capture the detail around the final conformation of tile B while the larger ones will ensure adequate roadmap coverage of the conformation space. The nodes around the final conformation of tile B will also be quite useful for locating local minima in which tile B may get trapped and never get matched with tile A properly.

The same goal can be achieved by manipulating the probability of accepting a configuration as a milestone which we will discuss now. Not every configuration sampled is accepted as a milestone. In particular, if we use random sampling, generating high and low energy conformations are equally likely. Hence, to bias our sampling process in order to generate tile B conformations in regions of low energy, a randomly generated tile conformation is accepted as a milestone with the following probability:

$$P(accepted) = \begin{cases} 0 & \text{if } E_{conform} > E_{max} \\ \frac{E_{max} - E_{conform}}{E_{max} - E_{min}} & \text{if } E_{min} \leq E_{conform} \leq E_{max} \\ 1 & \text{otherwise} \end{cases}$$

Here $E_{conform}$ is the energy of the tile conformation under consideration. E_{max} is the energy of tile B at any configuration lying on the boundary of the rectangular air bed. Note that the value of E_{max} remains constant throughout the whole experiment. E_{min} is the energy of tile B when it is bound with A in the desired manner.

We have two other special milestones- one corresponding to the initial conformation of tile B and the other corresponding to its final bound conformation. Let the set of milestones be denoted by V. In the next section we'll describe how to join a pair of nodes in V which will be our edge set $E^{'}$. This will define our roadmap $G(V, E^{'})$.

3.5.2 Roadmap Construction

Now that we have the milestones, we need a policy for joining them. Suppose we have sampled a total of n nodes. For every milestone m_i, we'll find the k nearest neighbors using Euclidean distance. Root Mean Square Distance(RMSD) is also popularly used. We set k to be equal to a multiple(say e) of the dimensionality d of the conformation space. Thus $k = ed$. We could have had the neighbor set N_i of m_i as all the milestones which are within Euclidean distance d_{cutoff} from m_i where d_{cutoff} is a user-defined parameter. So,

$$p \in N_i : ((x_p - x_{m_i})^2 + (y_p - y_{m_i})^2)^{\frac{1}{2}} \leq d_{cutoff}$$

Once we have found the neighbor set for all the milestones we sort the elements in each N_i in the increasing order of their Euclidean distance from $m_i, \forall i$ Now for each element $p_{ij} \in N_i$ we draw an edge between p_{ij} and m_{ij}, based on the local path planner method similar to the one described in [6].

Local Path Planner The modified local path planner for connecting two conformations is as follows: The planner connects two given milestones by a straight line in C space and determines whether this straight line path is energetically favorable. The path is tested by discretizing the line segment joining m_i and p_{ij} by a set of consecutive conformations For instance, a set R_{ij} of r random conformations which lie on the line segment $m_i p_{ij}$ are chosen. We accept the path $m_i p_{ij}$ as an edge in our roadmap if $\forall_{l=1}^{r}, E[\rho_{ijl}] \leq E_{max}$, where $\rho_{ijl} \in R_{ij}$. Thus, in algorithmic notation,

```
boolean localPathPlanner(node $v$, node $w$ )
    Define set $R$ as a set of $r$ random conformations which lie on the line segment $vw$
    for $l = 1 \ldots r$
        if $E[\rho] > E_{max}$, $\rho \in R$
```

```
        return false
    return true
```

Edge Generation Now we are ready to present the technique for drawing edges in our roadmap using the local path planner algorithm which is as follows:

```
void generateEdges(V)
    for $i = 1 \ldots n$
    let the $i^{th}$ milestone be $m_i$
    sort the remaining $n-1$ milestones in the increasing order of their
    Euclidean distance
    create the queue $N_i$
    while the $degree(m_i) \leq d$ and $N_i \neq \emptyset$
        $p_{ij} \leftarrow$ Dequeue($N_i$)
        if(localPathPlanner($m_i, p_{ij}$=true)
            $degree(m_i) = degree(m_i) + 1$
            $E = E \cup m_i p_{ij}$
```

Edge Weight Assignment Unlike traditional PRMs where the goal is to find whether a path exists between a start and end configuration, our goal is to find whether an energetically favorable path exists between the unbound state and the bound state of tile B. So we associate each edge $e \in E^{'}$ with a weight which is proportional to the potential energy of all the sampled conformations(from the local path planner) between the two milestones. For this we need to compute the probability of traversing the edge. Suppose $e \in E^{'}$ connects milestones m_i and m_j and the r consecutive conformations which discretize the line segment $m_i m_j$ be $s_1, s_2, \ldots, s_r$, then the probability of moving from one discretized conformation s_f the next one s_{f+1} is

$$P(s_f s_{f+1}) = \frac{e^{-(E(s_{f+1})-E(s_f))/kT}}{e^{-(E(s_{f+1})-E(s_f))/kT} + e^{-(E(s_{f-1})-E(s_f))/kT}}$$

Weight of the edge e is

$$w(e) = \sum_{i=1}^{r} - \log P(s_f s_{f+1})$$

Note that the sum of the probabilities for all the outgoing edges for a milestone may not sum to one. The remaining probability indicates the probability of staying in the same milestone(self-loop). An example of a PRM for a two dimensional configuration space is given in figure 12

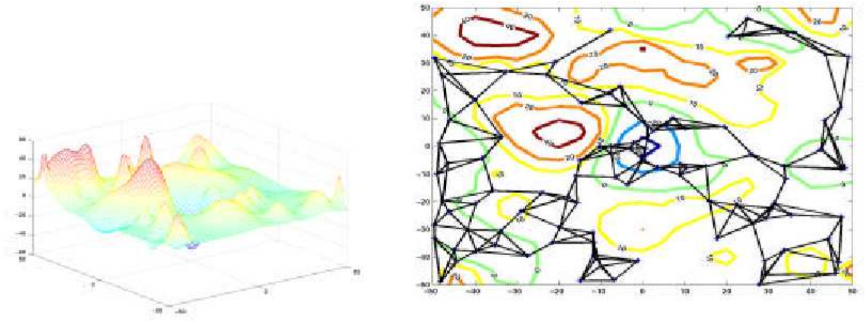

Figure 12: Probabilistic Roadmap(right) for a fictitous energy function(left) over two dimensional configuration space(Adopted from Apaydin 2003

3.6 PRM Search

Basic Query Now that we have created the weighted graph $G(V, E')$, we can query the PRM in order to find whether an energeically favorable path exists from the initial configuration of tile B to its final configuration. So for any randomly generated initial and goal configurations, we first check the reachability of the final configuration starting from the initial configuration. We can use depth first search for this purpose. In case a path exists, our next goal is to compute the minimum-weight path between the start and the goal configuration using Dijkstra's shortest path algorithm. It will give us the most energetically favorable path between the initial unbound state of tile B and its final bound state.

u best paths A tile does not always follow the most energetically favorable path. In that case, we need to find the average difficulty of reaching the final configuration t starting with the initial configuration s. However, finding all paths in a graph requires time exponential in the number of nodes So, instead we can compute the average energy lanscape with respect to different initial configuration of tile B and the same final configuration.. We can run Dijkstra's algorithm with each node in the roadmap as the start node, rank them in the increasing order of their total edge weights and consider the u most energetically favorable paths. where u is a user-defined parameter. Note that $u \leq n$. The goal configuration might not be reachable from every node in the roadmap and hence we take the u best paths, depending on the total weight of the path. We set a cutoff w_{cutoff}, such that if the total weight of the path is above the cutoff, the path is rejected. We decide on the value of u based on the value of number of paths accepted.

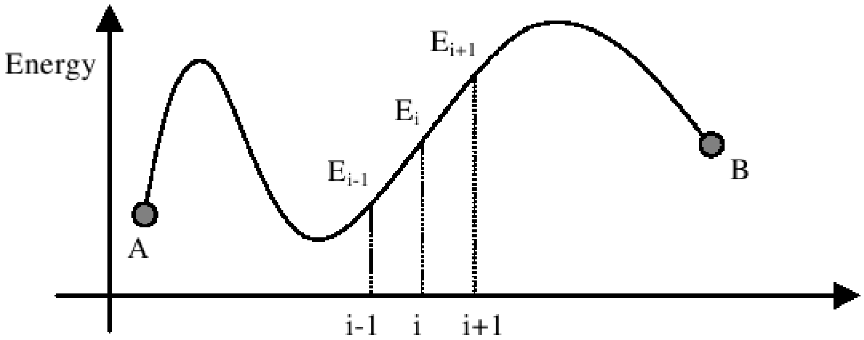

Figure 13: Energy Barrier around binding site for one dimensional configuration space

Query for different final configurations One of our main goals is to study how the potential energy landscape of tile B changes with the variation of relative tile configuration . This could be best studied if we consider several final configurations of the tile along with several initial configuration. We could have done this with a single PRM. However, our milestone generation procedure is biased by the final configuration of the tile. So we generate a new PRM for every final configuration of tile B and obtain the set of u most energetically favorable paths using the procedure discussed above.

Analysis If we plot the energy of the u best paths for a particular final configuration against the nodes in the path, we can have an idea how easy it is for tile B to move from an initial unbound configuration to the final bound configuration. We can also detect whether in general, an energy barrier exists between the initial and final configuration which decreases the probability of the assembly. Similarly, our study for different final configurations will demonstrate the typical behavior of tile B in its configuration space. Our conjecture is that there will be significantly deep energy barrier around the binding site, thus not only derceasing the chances of an assembly but also would make it significantly difficult for a bound tile to leave the binding site[Figure 13].

3.7 Characterization of a true binding site

A major problem with tile assembly by shape recognition is that in many cases we end up having partial matches and the incoming tile might not bind completely to the growing assembly. It might also end up in another

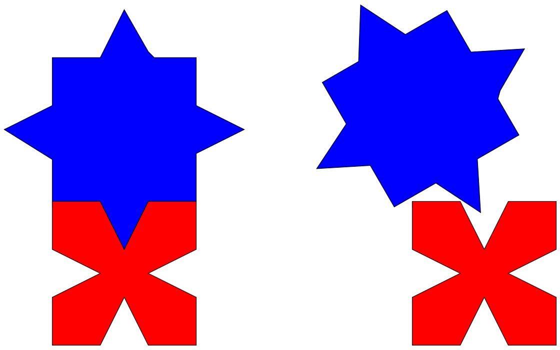

Figure 14: A true binding site corresponds to a complete match(left) while a partial binding site corresponds to a partial tile match(right)

binding site which has comparable energy to the true binding site and may also have partly complementary shape too[Figure 14]. We need to design the tile shapes and arrange the dipoles in such a manner such that the tiles are able to distinguish between true binding site and potential binding site. Two attributes that can be used to distinguish between true and potential binding site is the absolute energy of the bound configuration in each case and the average weight of the u best paths entering that configuration. If the difference between the absolute energy of a true binding site and a false binding site is large[Figure 15] then quite naturally the former will be preferred But we also believe that the relative difficulty of entering and leaving a true binding site will be more than a false binding site, indicating that the energy barrier in the former case is larger than the latter case. This can be obtained from computing the average path weight for each of the final configurations.

Ease of Tile Disintegration It is a well known fact that self-assembly is a reversible process but it depends on the energy of the system. In our case, as the energy barrier is high for a true binding site, as mentioned earlier it is as difficult to leave the binding site as it is to enter the binding site. This can be verified by swapping the original initial and final configuration Thus, the more difficult it is for tile B to get detached from tile A, the more is the binding affinity between the two tiles.

PRM path weight and rate of assembly The binding affinity and the energy barrier of the tile is also related to the rate of assembly . In other words, the tile is trapped in its binding site by the energy barrier We believe that the high difficulty weight for leaving the binding site indicates a very low rate of dissociation, which dominates the standard free energy of the reaction.

Figure 15: Distinguishing a true binding site from a false one

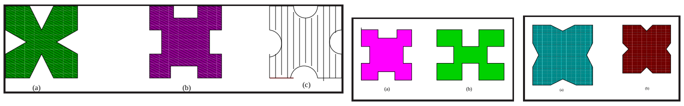

Figure 16: Comparison of shapes of different binding sites(left), comparison of depth of the binding sites(middle), comparison of the width of the binding sites(right)

3.8 Effect of tile shapes

Our main objective here was to study change in potential energy landscape of the moving tile with the change of the shape of the binding site. Precisely, our goal is to identify the role of binding site curvture, width and depth in distinguishing a true binding site from a false one. Ideally we do not want to get the tile B stuck in a pseudo binding site. Even if the tile tries to get attached to such a site, the energy difference between a true binding site and a false binding site has to be large enough so that the tile detaches itself from such an energetically unfavorable bound configuration and seeks the correct one. For instance, our conjecture is that increased curvature of the binding site would increase the ease of assembly while shallower binding site will have a smaller energy barrier than a deep one. Area of contact is also a contributing factor to the success rate of the assembly. and wide mouthed binding site is more easily accessible than a narrow ridge. We believe that our PRM analysis for different tile shapes will allow us to understand the correlation between binding site curvature, width or depth and the binding affinity or ease of assembly.

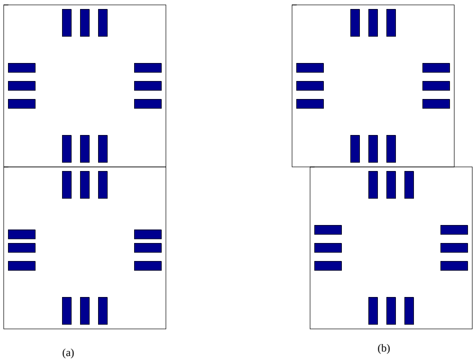

Figure 17: Complete Match(left) and partial match(right) for self assembly without shape recognition

3.9 More Applications

The PRM is not only useful for understanding the variation of the potential energy landscape with the variation of tile shapes, but also when the assembly happens not by shape recognition but by other interaction. For instance, if we had just square tiles, and the assembly was taking place only by the pairing up of the complementary dipoles[Figure 17], even then correlation between the potential energy of the tile and the dipole placement would allow us to design better tiles and hence improve the sucess of tile-assembly. We could conduct the same set of experiments as discussed before- query the existence of an energetically favorable path from inital unbound configuration to final bound configuration and study the height and width of the energy barrier around the true binding site and compare it with pseudo binding site and thus gain some insight into the problem of optimal placement of the dipoles for maximum yield from our self-assembly experiments.

4 Future Work

The most immediate future work is the actual implementation of our basic probabilistic roadmap technique for a two tile assembly using shape recognition and study the effect of various shapes on the potential energy landscape. The basic implementation calls for a fixed 2D tile and a moving 2D tile. We would also like to study the effect of removing the positional constraint on the other tile and later learn about the typical behavior of a tile in a multi-tile system. This would require generating the probabilistic roadmap in more than three dimensions. We would also like to modify the basic software so that we can study the energy barriers involved in self-assembly using only magnetic or capillary interaction. Finally we would like to extend the sofware so as to study the behavior of three dimensional tiles by generating its energy contour in at least a six dimensional configuration space.

5 Conclusion

In this project we studied the dynamics of self assembly using a motion planning algorithm. Our technique uses magnetic potential of the interacting tiles to compute the most energetically favorable path between any given initial and final tile configuration. We use probabilistic roadmap to sample the distribution of possible paths to a given goal configuration and compute an energy based ”difficulty weight”(energy barrier) for each path. which is indicative of the relative difficulty of entering or leaving a certain goal configuration. This approach captures the whole energy landscape, in particular the area around the binding site and can be used to study the effect of tile shapes in distinguising a true binding site from a false one. Our hope is an actual implementation of our probabilistic roadmap will help us verify our hypotheses and provide us with a better tile design that would improve our success rate in self-assembly experiments

References

[1] Mila Boncheva, Derek A. Bruzewicz, and George M. Whitesides. Millimeter-scale self-assembly and it applicationss. *Pure Appl. Chem.*, 75(5):621–630, 2003.

[2] Rothemund PW. Using lateral capillary forces to compute by self-assembly. *Proc Natl Acad Sci U S A*, 97(3):984–989, 2000.

[3] Zeng T, Claus R, Liu Y, Zhang F, Du W, and Cooper KL. Piezoelectric ultrathin polymer films synthesized by electrostatic self-assembly processing. *Smart Mater. Struc.*, 9:801–804, 2000.

[4] Boncheva M, Andreev S, Mahadevan L, Winkelman A, Reichman D, and Prentiss M. Magnetic self-assembly of three-dimensional surfaces from planar sheets. *PNAS*, 102(11):3924–3929, 2005.

[5] Fang J, Liang S, Wang K, Xiong X, and Bohringer KF. Self-assembly of flat micro components by capillary forces and shape recognition. In *FNANO 05*.

[6] Kavraki LE, Svetska P, Latombe JC, and Overmars M. Probabilistic roadmaps for path planning in high dimensional configuration spaces. *IEEE Trans. Rob. and Autom*, 12(4):566–580, 1996.

[7] Apaydin MS, Singh Ap, Brutlag DL, and Latombe JC. Capturing molecular energy landscapes with probabilistic conformational roadmaps. Int. Conf. on Robotics and Automatons, 2001.

[8] L Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 226(11):1021–1024, nov 1994.

[9] J-C Latombe. *Robot Motion Planning*. Boston: Kluwer, 1991.

[10] Lozano-Pérez T and Wesley M. A. An algorithm for planning collision-free paths among polyhedral obstacles. *Comm. ACM*, 22(10):560–570, 1979.

[11] Dúnliang C Ó and Yap C. K. Aretraction method for planning the motion of a disc. *J. Algorithms*, 6:104–111, 1982.

[12] Canny J F. *The Complexity of Robot Motion Planning*. MIT Press Cambridge MA, 1988.

[13] Singh A P, Latombe JC, and Brutlag DL. A Motion Planning Approach to Flexible Ligand binding. In *Proc* 7^{th} *Int. Conf. on Intel. Sys. for Mol. Bio.* AAAI Press, Menlo Park, CA.

[14] Anderson H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys*, 72:2384–93, 1980.

[15] Rubinstein R Y. *Simulation and Monte Carlo Methods*. 1981.