Development of an Agent based Model illustrating the usage of Deferred Acceptance Algorithm in the Admission Process

Abhishek Bhatia[1,2*], Chetan Sharma[1,3], and Rinkaj Goyal[1,4]

[1]**USICT, GGS Indraprastha University, New Delhi**
[2]***Corresponding Author; Email: abhigenie92@gmail.com**
[3]**Email: cs15june@gmail.com**
[4]**Email: rinkajgoyal@gmail.com**

ABSTRACT

Agent-Based Modelling framework successfully models real life problems that support simulation with diverse strategies and mechanisms devoid of the restrictions set by mathematical tractability. Union of game theory and agent based modelling has elucidated the dynamics of different social and economic scenarios. In this study, we present our efforts to develop an agent based model through embracing a customized version of the deferred acceptance algorithm. This study considers two widely adopted admission process scenarios i.e. Partially and Fully centralized as a case study, wherein a University acts as a nodal bureau and admits students to affiliated colleges. In this paper, an agent based model has been developed in the Netlogo simulation environment, which advocates fully centralized procedure and simulates Deferred Acceptance Algorithm. The simulation results in a strategy-proof,optimum and stable allocation of available seats in the admission process.

Keywords: Agent based modelling, Game theory, Defrerred Acceptance Algorithm, Netlogo

INTRODUCTION

Agent based modeling (ABM) provide computational techniques to design, investigate and experiment with agents manifesting an interaction within the situated environment (Gilbert, 2008; Bandini et al., 2009). These agents may be institutions, organizations, human beings or any other object that seeks its individual objectives (Railsback and Grimm, 2011).

ABM embraces basic features of complex systems such as self-organization, chaos and adaptation, which otherwise could have very difficult to achieve using mathematical formulation. Such modeling adapts to simulate complex systems in scenarios like economics, integrative biology, social network analysis and urban planning among others (Taghawi-Nejad, 2013; Holcombe et al., 2012; El-Sayed et al., 2012; Schwarz et al., 2012). More complex phenomena can be modelled by integrating ABM with other evolutionary algorithms (Bonabeau, 2002; Bouarfa et al., 2013).

In this study, we employ the advantages of ABM and cooperative game theory to solve the over-admittance problem in the admission of undergraduate courses to different colleges facilitated by a nodal agency (University). To demonstrate the effectiveness of the developed model, we consider two widely used admission scenarios, as discussed in the succeeding sections;

Scenario 1: Partially centralized admissions

In partially centralized admissions, university performs a minimum role by making the whole process decentralized. All colleges affiliated with the university decide and declare their own cutoff list of marks for admission each academic program. The cutoff list depends on the individual colleges past admissions related experience and present years high school results announced. Eventually, the aspirants who get marks beyond the cutoff of a college can claim a seat within a given stipulated period. This procedure repeats until there are no unclaimed seats.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.825v1 | CC-BY 4.0 Open Access | rec: 11 Feb 2015, publ: 11 Feb 2015

Following is the brief outline of the partially centralized admission process (Fig. 1). The factors

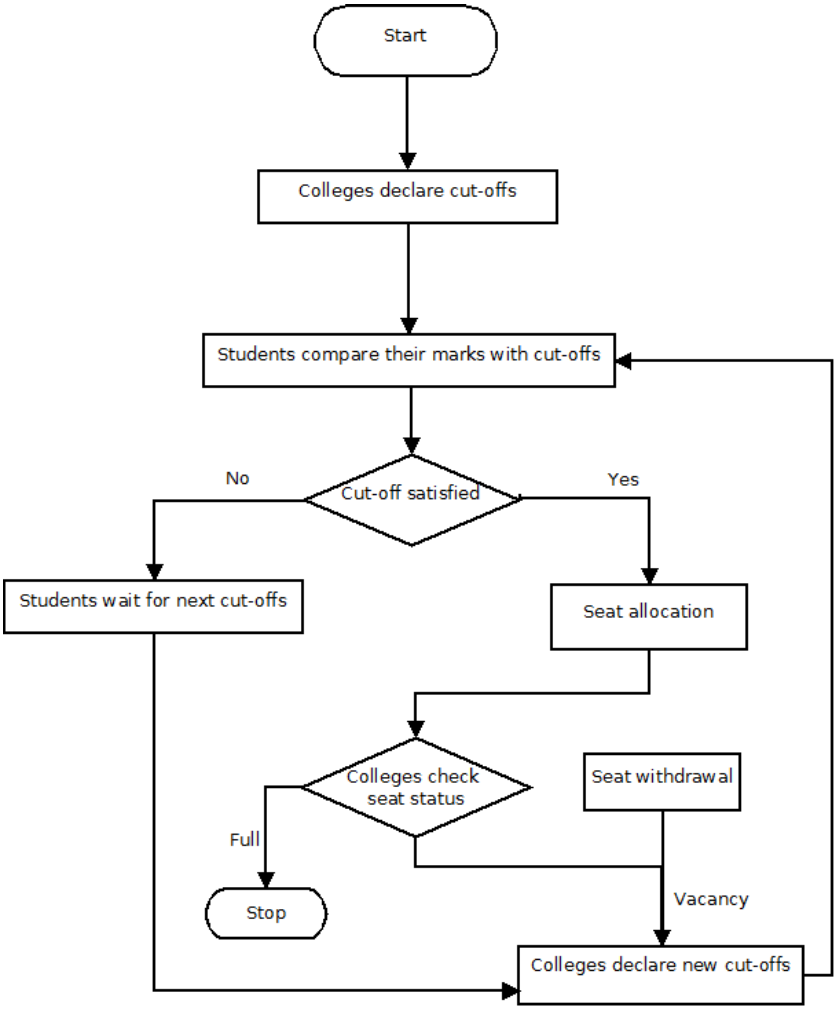

Figure 1. Partially centralized admission process.

influencing the admission process indicated above are;

1. Optimum cut-off formulated by an individual college based on their prior experience
2. Time frame to for the candidate to claim or vacate the seat

Now, this involves uncertainty in the whole of the decision making process on the part of college authorities. The college administration expects the maximum admissions meeting the initial cut-off declared. Initially, college over-expect and advertise a high first cut-off, which gradually keeps on sliding in the subsequent cutoff lists.

This process advances smoothly for the first two or three cut-offs. Subsequently, even the slightest drop in cutoff results in tipping phenomenon. This leads to the sudden escalation in the number of claims.

Since colleges cannot turn down the application of any of the students who meet the eligibility criteria, they have to admit all those applicants. This leads to the surplus admittance of students in many of the colleges, and sometimes the number of admits are as high as twice the total number of seats, resulting in an immense pressure on the admitting colleges.

This problem is a representative case of stable matching in the resource allocation wherein resources are the seats available in a college and students seeking admission are the players. A stable matching does not yield a profit to any player in further trade and eliminates justified envy (Chen and Sönmez, 2006). This refers to the scenario when a college either has a vacant seat or admits an applicant with a lower priority.

Alternatively, in aforementioned problem, a matching is unstable if there exists a student-college pair (i, c), where c is the preferred choice of other student having higher precedence.

Scenario 2: Fully centralized admissions

Following is the flow of information in the Fully centralized admissions (Fig. 2);

1. All students apply for the course of their first choice. Each college tentatively assigns its seats to its applicants one at a time on the basis of their performance in High school result. A student older in age will get preference in the determination of the inter-se-merit. Other students will be rejected.
2. Student rejected in the previous step further apply for their next choice if available. Each college takes such applications into consideration together with new applicants and tentatively assigns seats by following the procedure mentioned in the earlier step.
3. Remaining applicants are rejected.

MODEL DEVELOPMENT

The model developed here adapts the work of Gale and Shapley known deferred acceptance algorithm. Shapley and Roth won Nobel prize of economics in 2012 for the applicability of their work in the practical solution of economic problems (Roth, 2008). The simulation results empirically demonstrate the suitability of deferred acceptance algorithm to solve the problems faced by the partially centralized admission process. The figure below shows the key steps of the proposed algorithm which has been implemented in Netlogo agent based simulation environment (Fig. 3).

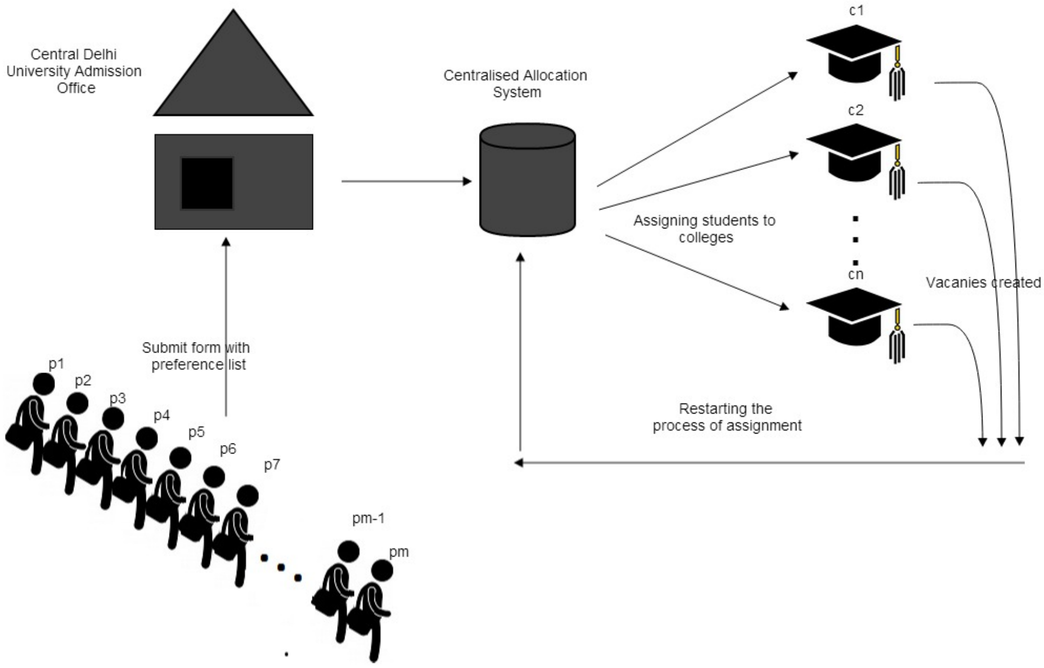

Figure 2. Fully centralized admission process.

The algorithm terminates when each student receives his final assignment. In case of the withdrawal of allotted seats, the algorithm starts from reassigning seats to the remaining aspirants. The centralized system needs a single centralized application form across all the colleges administered by the university.

In this form students will rank courses in the respective colleges, according to their preferences. Such a centralized system demands ensuing two mechanisms in place;

1. A mechanism for eliciting the preferences of the students
2. A mechanism to aggregate these elicited preferences and allocate seats to the respective colleges.

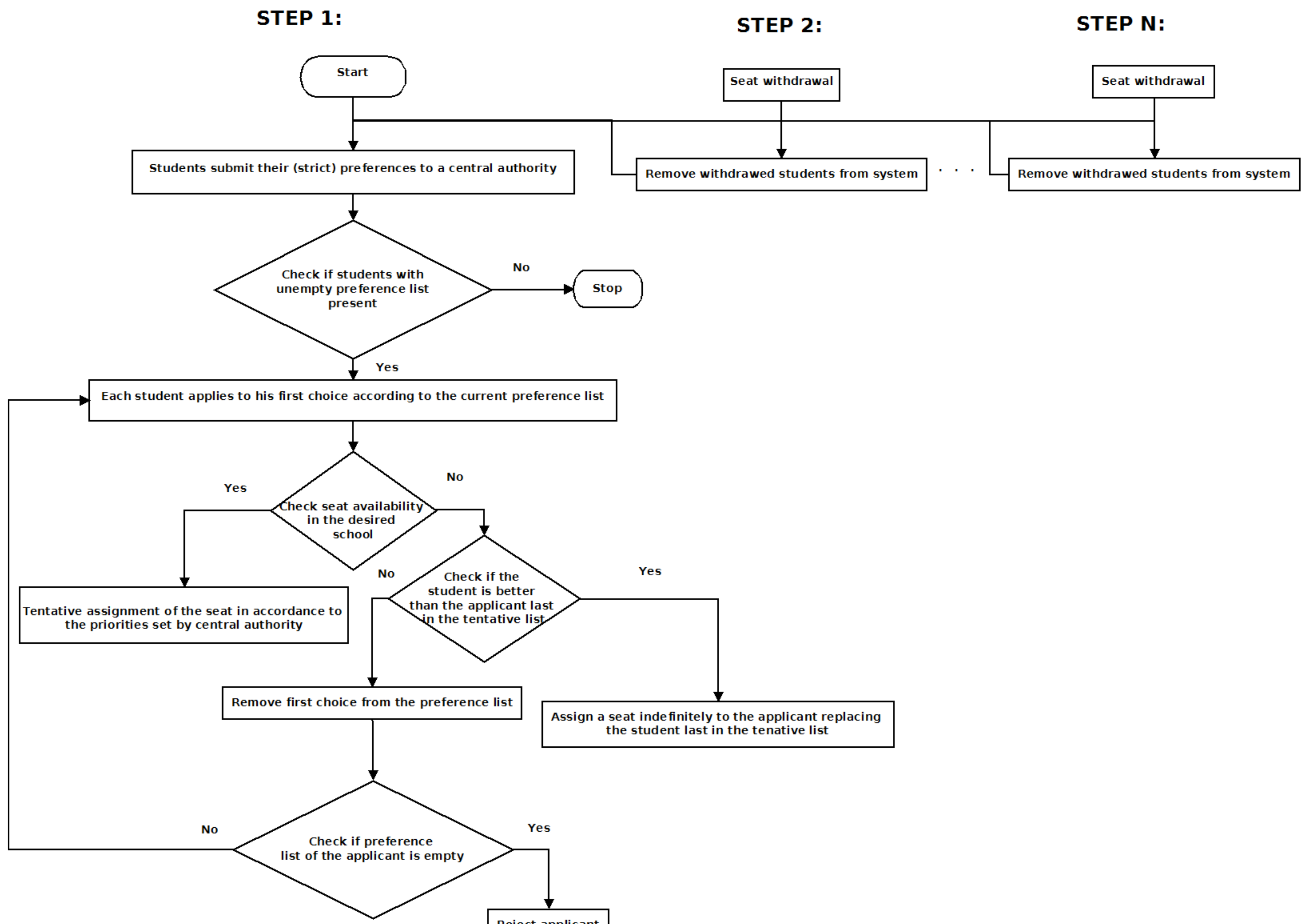

Figure 3. Proposed centralized admission process.

STABILITY ANALYSIS OF PROPOSED MODEL

This matching problem with aggregating stated preferences becomes a non-cooperative game between the students. Their payoff is the allocation of desired seat and strategy is preferences submitted. This procedure gives students the incentives to reveal their true preferences and dominant strategy for each student to state his honest preferences.

No matter what preferences other students may state, a student who falsifies his preferences can achieve no better outcome than if he had stated them correctly (Chen and Sönmez, 2006; Sönmez, 1997).

The discussion below illustrates that this deferred mechanism result in a strong Nash equilibrium. There do not exist any incentive for any student to misrepresent his true preferences and end up with a better college.

In a strong Nash-equilibrium profile, There does not exist a coalition of members, which can deviate by changing strategies to benefit all of its members, provided non members keep their strategies unchanged Consider a situation where a set of three students labelled S_1, S_2, S_3, $S_4$and S_5 with marks 71, 72, 70 ,75 and 76 respectively, is to be matched against a set of colleges namely C_1, C_2, C_3 (each with maximum seat limit 2 in C_1 and 1 in C_2 and C_3), assuming, a student is not indifferent to two colleges that is student has a strict preference relation over any two colleges and does not prefer staying unmatched therefore, student accepts the worst possible college in his list but not stay being unallocated. Let the true preferences for students be given as follows:

Students(S_i)	**Percentage**	**Preferences over colleges(C_i) in order (>)**		
S_2	72	C_1	C_3	C_2
S_3	70	C_2	C_1	C_3
S_4	75	C_2	C_1	C_3
S_5	76	C_1	C_3	C_2

Case I: Students quote their true preference lists:
Seat distribution in each college after allocation procedure has completed.

C_1*(2 seats)*	**C_2** *(1 seat)*	**C_3***(1 seat)*
S_5	S_4	S_1
S_2		

Case II: Student S_2 misrepresents his true preferences:

1. S_2: $C_3 > C_1 > C_2$

Seat distribution in this case in each college after allocation has completed.

C_1*(2 seats)*	**C_2** *(1 seat)*	**C_3***(1 seat)*
S_5	S_4	S_2
S_1		

1. S_2: $C_2 > C_1 > C_3$

For this case the seats in each college are assigned after procedure has completed.

C_1*(2 seats)*	**C_2** *(1 seat)*	**C_3***(1 seat)*
S_5	S_4	S_2
S_1		

Student S_2cannot do better than the 1^{st} best college in his true preference list and in case if he deviates from his true preferences he might end up being matched to worse choices.
Hence, the above discussed student proposal procedure is a strong Nash Equilibrium and entirely strategy proof where there can be no incentive for any student to manipulate his true preferences.

Fairness for students

Now consider students s1 and s2, their scores and their respective true preference lists.
s1: c1 > c2 > c3 ;; s1: 71
s2: c1 > c3 > c2 ;; s2: 72

As per the college preference lists, we can figure out that both the students rate c1 as their first choice irrespective of the marks they score. For the model to be stable, s2 must be preferred over s1, i.e. s1 must not be assigned a seat in c1 before s2, under any circumstance whatsoever.

Under the one-sided interpretation explained above, stability embodies a notion of fairness: a student should not envy another school over her assignment, and have a higher claim to that school. A student with a better score and lesser age compared to other has a higher priority for a college, thereby endorsing the concept of serial dictatorship (Batty, 2007).

Stability and Optimality for Colleges

The problem posed in this study is closely related to the well-known school choice problem (Garavaglia et al., 2013). The following section explains how it achieves the stability and optimality for colleges along with fairness for students.

The above mentioned matching method yields the student- optimal stable assignment of students to colleges, and gives no student any incentive to misrepresent his preferences. Since it is a dominant strategy for each student to reveal his true preferences, the only potential source of distortion of the procedure now lies in the priorities of the colleges over the applying students and stating their true capacities.

There exists no stable matching mechanism which makes it a dominant strategy for all colleges to state their true preferences, although as proved above the deferred acceptance algorithm makes it a dominant strategy for all students to state their true preferences (Olsson, 1965; Castle and Crooks, 2006). There are no stable procedures that are strategic proof for both sides.

RESULTS AND DISCUSSION

This study is simulated in Netlogo environment (Tisue and Wilensky, 2004). The strategy proof allocation model allows agents (students) to list their true preferences without any manipulations to obtain the best possible choice from their preference list. Stochasticity in the simulation of model is introduced by initializing preference lists of students and assigning random values of percentages and ages (to break any ties). Following Observations were made

1. Cutoffs of colleges versus (their) Reputation
2. Number of students versus Preference acquired
3. Cutoff s of colleges versus Time Percentage of students versus Preference acquired by students

The reputation of a college is defined as follows; A college which is ranked higher among the preference lists submitted by the students will have more reputation consequently. The reputation is of a college can be calculated as:

$$R(\ c) = \sum_{j=0}^{j=n} e^{\ ((numberofcolleges)\ -\ j)} \tag{1}$$

Where R (c) denotes the reputation of college c, $(numberofcolleges)$ signify the number of available colleges for an assignment and i give the rank of current college c in the preference of present student. The summation sums across all the students for each college. n denotes the total number of applicant students. The instantaneous values of the number of students presently allocated, the number of students presently unallocated students with a remaining preference list are also being analyzed.

Figure 4 shows that the number of students of each order of the preference acquired college at the termination of the algorithm. Figure 5 depicts the allocation state after vacation of a given number of students (n=5, here).

As expected, to presently apply, students get into more preferred colleges depicted by a rightward shift in the histogram shown in Figure 5.

Figure 6 demonstrates that students with higher percentages manage to make it to more preferred college in their true preference lists, thereby satisfying the concept of fairness and serial dictatorship as discussed earlier.

After the execution of an iteration of the algorithm, each colleges cutoff is the least percentage among the presently admitted students. Student age decides inter-se-merit. Figure 7 represents the variance in

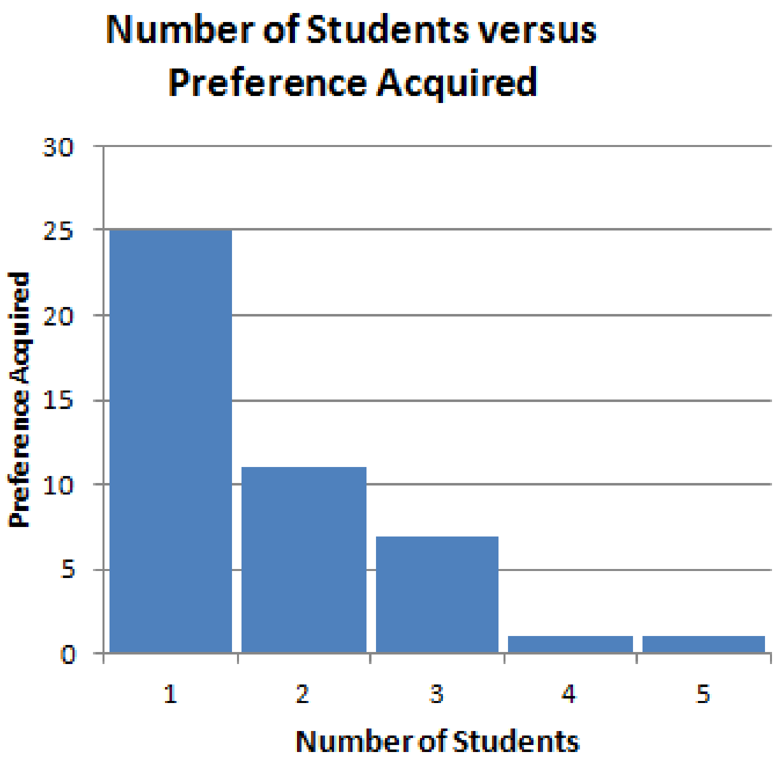

Figure 4. Number of students after the termination of algorithm .

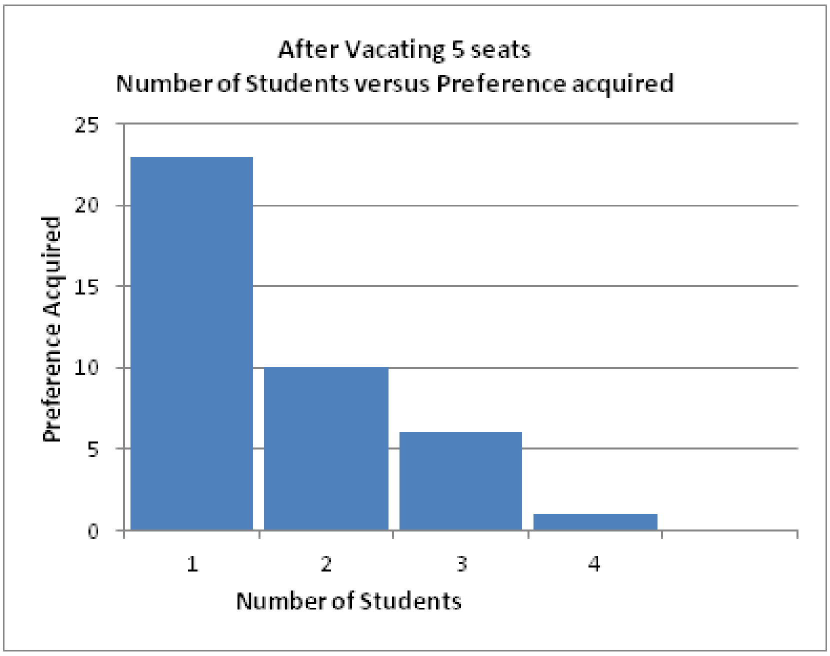

Figure 5. Allocation after the withdrawal of students .

the cutoff in the succeeding iteration of the procedure. The cutoffs decrease when students withdraw allotted seats as anticipated.

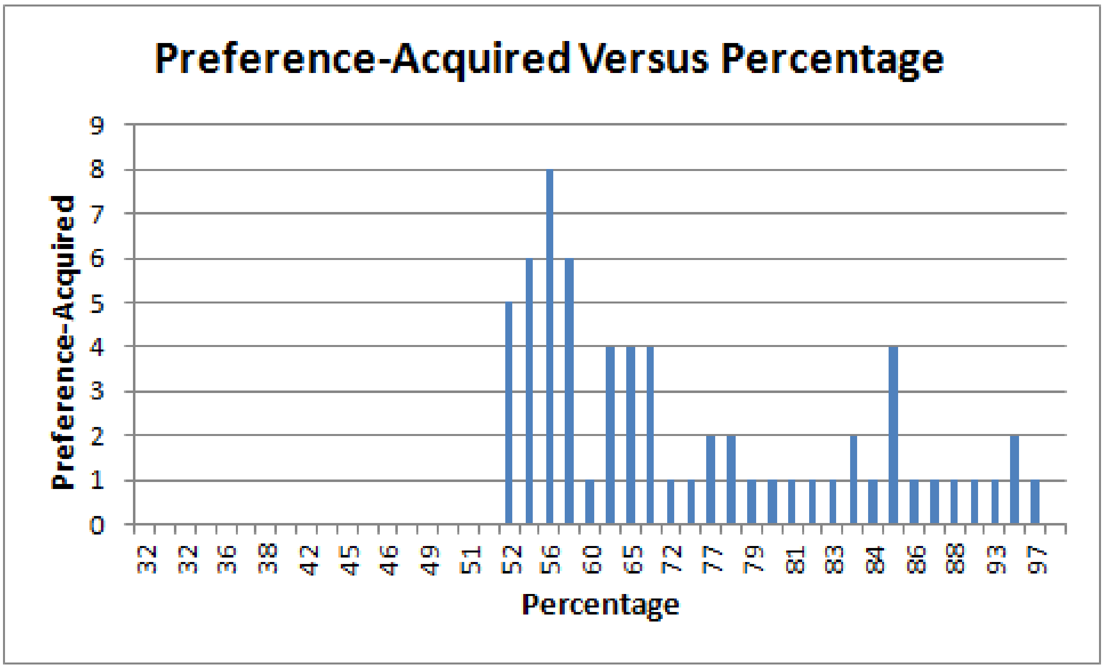

Figure 6. Illustration of serial dictatorship.

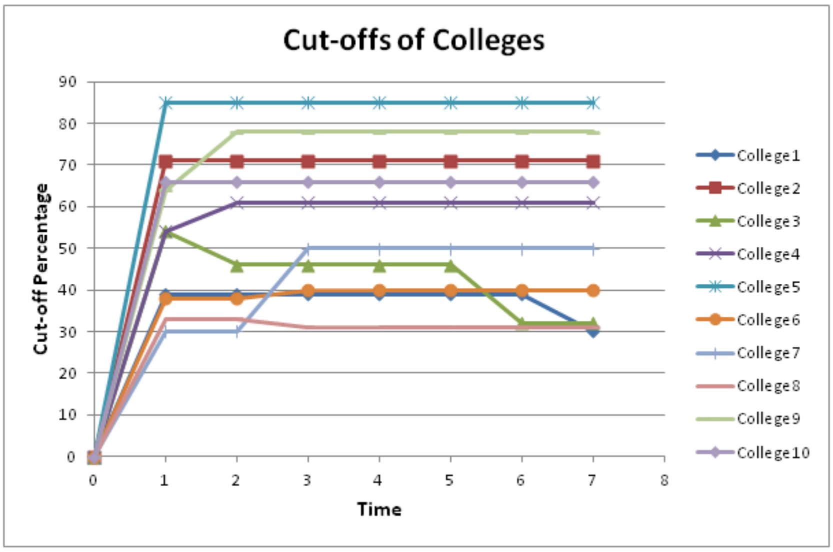

Figure 7. College wise variation in cut-off marks.

Figure 8 illustrates the relation between the present cutoff of colleges and their respective reputation calculated according to equation 1.

An interesting observation of this model exhibits that the most favorite colleges might not have the highest cutoff. Since, colleges decide these cutoffs wholly based on their experience, including the metric result of the present year. Colleges often exaggerate and over-estimate the true cutoff values. Whereas, the method we have used here to estimate reputation of a college takes into account the position where all the students place this college in their true preference-list and then formulate cutoff accordingly.

Now even if colleges somehow predict the preferences of students with the data of present marks of student applicants. There is no way to know the spread of the percentages of students for a college.

This situation corresponds to the theory of deterministic chaos, whereby small differences in the initial conditions leave larger differences in the outcome. Thus, subtle incorrect predictions about the

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.825v1 | CC-BY 4.0 Open Access | rec: 11 Feb 2015, publ: 11 Feb 2015

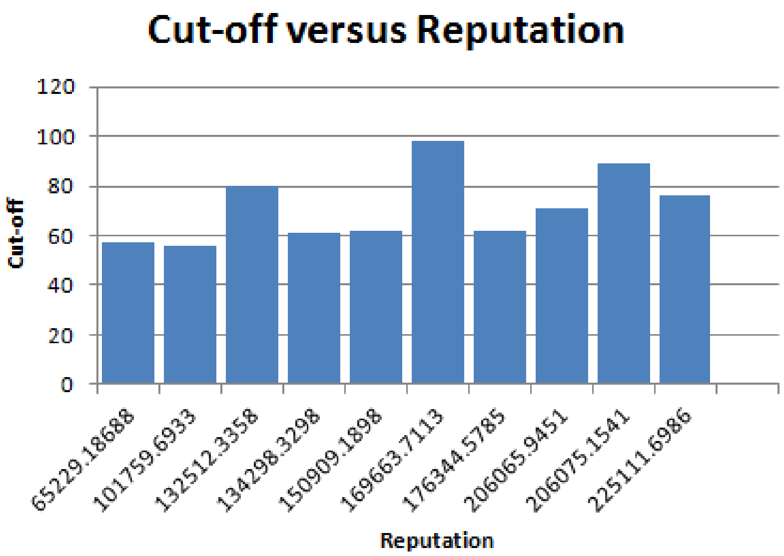

Figure 8. Cutoff v/s Reputation.

cutoff may result in chaotic unpredicted allocation. Conclusion The model developed in this using using fully centralized admission process results in seats allocation with no unstable student-college pair, thereby provides a stable matching of students to colleges. Besides tackling the problem of over-admittance, this model eradicates the major issue of formulation of accurate cutoffs by colleges.

Given the huge number of affiliating colleges in a university with their numerous respective courses, the size of the preference list allowed for a student for application must be limited due to the computational time complexity of the algorithm. At the same time the preference list must should provide student a good chance of representing most of his preferred courses in the desired colleges. Since it is not a dominant strategy for colleges to reveal their true capacity, therefore, to overcome this drawback for matching, each colleges maximum seat availability must be verified by the nodal agency i.e. University.

REFERENCES

Bandini, S., Manzoni, S., and Vizzari, G. (2009). Agent based modeling and simulation: An informatics perspective. *Journal of Artificial Societies & Social Simulation*, 12(4).

Batty, M. (2007). *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 3):7280–7287.

Bouarfa, S., Blom, H. A., Curran, R., and Everdij, M. H. (2013). Agent-based modeling and simulation of emergent behavior in air transportation. *Complex Adaptive Systems Modeling*, 1(1):1–26.

Castle, C. J. and Crooks, A. T. (2006). Principles and concepts of agent-based modelling for developing geospatial simulations.

Chen, Y. and Sönmez, T. (2006). School choice: an experimental study. *Journal of Economic theory*, 127(1):202–231.

El-Sayed, A. M., Scarborough, P., Seemann, L., and Galea, S. (2012). Social network analysis and agent-based modeling in social epidemiology. *Epidemiologic Perspectives & Innovations*, 9(1):1.

Garavaglia, C., Malerba, F., Orsenigo, L., and Pezzoni, M. (2013). A simulation model of the evolution of the pharmaceutical industry: A history-friendly model. *Journal of Artificial Societies and Social Simulation*, 16(4):5.

Gilbert, N. (2008). *Agent-based models*. Number 153. Sage.

Holcombe, M., Adra, S., Bicak, M., Chin, S., Coakley, S., Graham, A. I., Green, J., Greenough, C.,

Jackson, D., Kiran, M., et al. (2012). Modelling complex biological systems using an agent-based approach. *Integrative Biology*, 4(1):53–64.
Olsson, G. (1965). Distance and human interaction. a migration study. *Geografiska Annaler. Series B, Human Geography*, 47(1):3–43.
Railsback, S. F. and Grimm, V. (2011). *Agent-based and individual-based modeling: a practical introduction*. Princeton University Press.
Roth, A. E. (2008). Deferred acceptance algorithms: History, theory, practice, and open questions. *International Journal of Game Theory*, 36(3-4):537–569.
Schwarz, N., Kahlenberg, D., Haase, D., and Seppelt, R. (2012). Abmland-a tool for agent-based model development on urban land use change. *Journal of Artificial Societies & Social Simulation*, 15(2).
Sönmez, T. (1997). Manipulation via capacities in two-sided matching markets. *Journal of Economic Theory*, 77(1):197–204.
Taghawi-Nejad, D. (2013). Modelling the economy as an agent-based process: Abce, a modelling platform and formal language for ace. *Journal of Artificial Societies & Social Simulation*, 16(3).
Tisue, S. and Wilensky, U. (2004). Netlogo: A simple environment for modeling complexity. In *in International Conference on Complex Systems*, pages 16–21.

APPENDIX A

Model Functioning Guidelines

Entities and State Variables

Entities	Variables/Characteristics	Description
Student	percent	Aggregate percentage of marks in XII sta dard
	age	Age of candidate
	admit	Boolean variable set to false if the student presently not admitted and true otherwise
	pref-list	The preference of a student is this list, whi is randomized among the colleges present the start.
College	*max-seats*	Designates the maximum seats present in t college, which randomized at the start
	seats-count	Contains presently allocated seats in the co lege
	coll-admits	Stores the sorted list of currently admitte students to the college in the descending o der by percentage and age
	present-cutoff	To store the cutoff of college, which is t least percentage of the admitted students the moment.
	reputation	Decided according the preference list by t students

Initialization

User select number of colleges and students. Each student is unassigned at the start with randomly generated preference list and percentage. Every college is presently un-allocated with the reputation based on the preference of the students. The students have random color at the beginning, which later changes to the vividness of the college to which they are assigned upon allocation or turns to gray if they are unassigned. The algorithm is carried out by the sub-model " **allocate**" assigning students to their preferred colleges.

Sub-models

The sub-model " **allocate**" is the implementation of customized Gale-Shapley procedure. It runs for all unassigned students with a nonempty preference list. The preference list keeps getting reduced as the

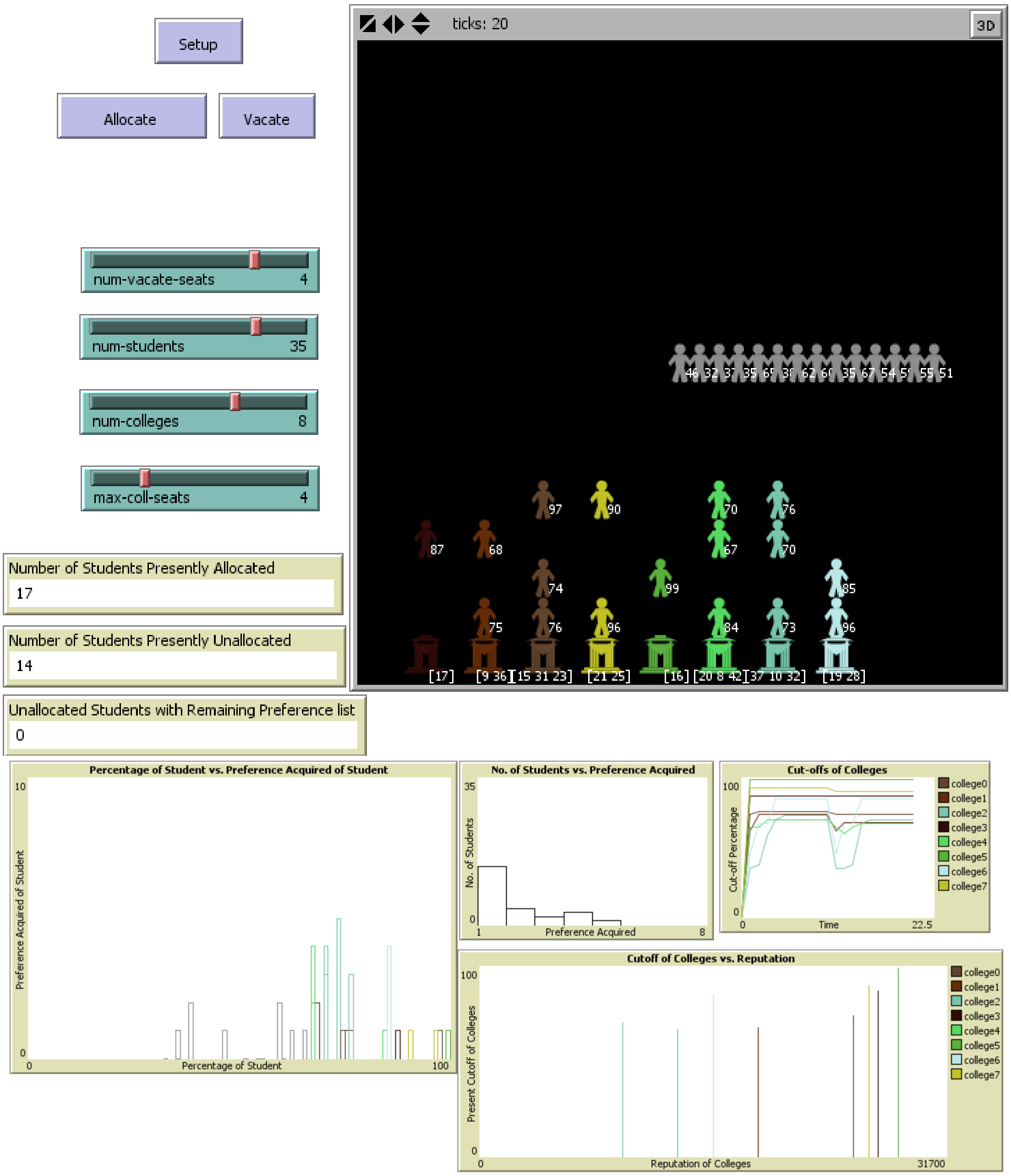

Figure 9. Snapshot of the model developed in Netlogo.

students are rejected by the respective colleges of their preference, or remains constant upon admittance to a particular college until the present cutoff (given by the variable present-cutoff) rises above his percentage. If this happens then the student is replaced by student with higher percentages. In case of same percentages age factor (age) is used to break ties. The sub-model " **vacate**" is used to vacate the desired number of assigned students from their admitted colleges, which in turn removes all presently allocated students after removing the specified number of admitted students. This model consequently calls the " **allocate**" sub-model after vacation to restart the process allocation of students to colleges.

Process Overview and Scheduling

When the model starts the students and colleges are created and configured with the characteristics mentioned in the Entities and State Variables section. The following steps are executed, in this order, once per time step:

- If all students " ***pref-list***" is empty or all students have been admitted (that is, admit is set true for the particular student) to one of his desired colleges then the model stops and the allocation process is complete.
- The sub-model "allocate" executes for all those applicants whose "pref-list"is still nonempty and their admit is set to false. The order in which the students apply to the their respective preferred colleges is shuffled at each time step.
- As the sub-model " **allocate**" executes, the admit variable of students is updated and *seats-count*, *coll-admits*, *present-cutoff* of colleges also changes consequently.
- The plots for variation of various state variables are updated each time step with the execution of " **allocate**".
- For removing a desired number of assigned students from their allocated colleges after the process of allocation we use the sub-model " **vacate**".

How to use this model

- Press SETUP to create the colleges and students, thereby assigning each student randomly a percentage and age along with a college "preference list" for admission or allocation. A visualization of each colleges current allocation of students will be initiated.
- Press ALLOCATE to run the model to allocate students to their preferred colleges. Press VACATE to the number of students specified by the variable *num-vacate-seats*
- The NUM-STUDENTS slider specifies the number students applying for admission in the colleges available.
- The NUM-VACATE-SEATS slider decides how many presently allocated students you want to remove from their assigned colleges.
- The NUM-COLLEGES slider is used to specify the number of colleges for students allocation
- The MAX-COLL-SEATS slider is used to limit the maximum number seats of colleges for admittance of students.