Understanding the Screening Process for New Molecules

James Hook [1], Kamil Kulesza [2] Piotr Morawiecki [2] and R. Eddie Wilson [3]†

[1] *Department of Mathematical Science, University of Bath, UK*
[2] *Centre for Industrial Applications of Mathematics and Systems Engineering, Polish Academy of Sciences, Warsaw*
[3] *Department of Engineering Mathematics, University of Bristol, UK*

(*Communicated to* MIIR *on 22 April 2021*)

Study Group: ESGI 138. 16–20 July 2018, Bath.

Communicated by: Hilary Ockendon

Industrial Partner: Syngenta

Presenter: Kim Travis

Team Members: James Hook, Bath; Kamil Kulesza, Warsaw; Piotr Morawiecki, Warsaw; Eddie Wilson, Bristol

Industrial Sector: Agriculture/Fisheries

Tools: Optimisation, Monte Carlo, Prediction markets, Markov decision theory

Key Words: screening, simulation, chemical synthesis, prediction

MSC2020 Codes: 90, 91, 92

† Corresponding Author: r.e.Wilson@bristol.ac.uk

Report Authors and Contributors

James Hook (Department of Mathematical Science, University of Bath)

Kamil Kulesza (Centre for Industrial Applications of Mathematics and System Engineering)

Piotr Morawiecki (Centre for Industrial Applications of Mathematics and System Engineering)

R. Eddie Wilson (Department of Engineering Mathematics, University of Bristol)

Facilitator:
Contributors:

Contents

1 Executive Summary

1.1 Problem Background

Globally there is a huge market for herbicides. Syngenta and its competitors spend large amounts of money in trying to develop new herbicides which are highly effective (kills the plants that the farmer wants to get rid of), selective (does not harm the crops that the farmer is trying to protect), safe (not harmful to humans or the environment) and cheap to produce.

Every year Syngenta chemists develop around one thousand new compounds that could be used as herbicides. Accurately evaluating the efficacy, selectivity, safety and ease of production of each one of these compounds would be extremely expensive. Initially synthesizing compounds even in tiny quantities requires a lot of work. Different compounds will then require different levels of dosage, may target certain species of plants better than others, may work more effectively in conjugation with different solvents etc. A huge amount of time and money could be spent optimizing the application of a particular compound, which then turns out to be very poor compared to existing herbicide products on the market.

In order to attempt to select only the best performing compounds in a cost effective way, Syngenta use a screening cascade. This process begins by testing all of the candidate compounds in laboratory experiments, referred to as assays, to see if they can (e.g.) target a particular enzyme, or penetrate a leaf. These lab experiments only require a tiny quantity of each compound to be synthesized. Compounds which perform poorly in these first experiments are discarded from the trial. We call this the first screen.

Compounds which pass this first round are then passed to a second round in which a small amount of the compound is applied to several small pots containing a few different species of plant. Compounds which perform poorly in this second experiment are discarded from the trial process. We call this the second screen.

In subsequent screens larger quantities of the compounds are used in the experiments, which makes them a lot more expensive. By using more compounds the candidate herbicides can be tested on a larger range or species, at different dosages, in conjunction with various different conditions, and with less sampling error.

As the screening process progresses, the screens become steadily more rigorous and consequently more expensive. At the same time the number of compounds remaining in the trial goes down. Finally a small number of compounds are taken to the final level, called the field trial. In this trial the compounds are applied outdoors in the way they would be
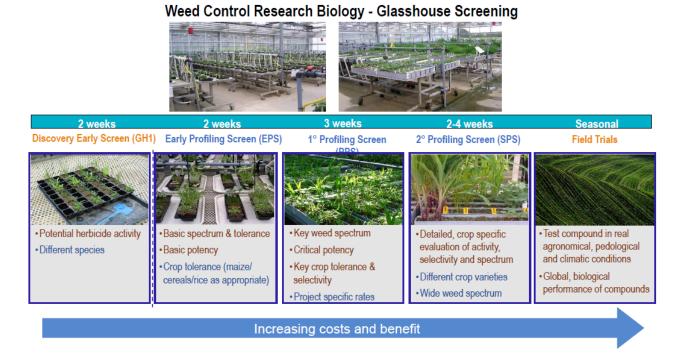
Figure 1: Five stages of the screening cascade.

used if they were eventually developed into a commercial product. See Figure 1.

1.2   Problem Statement

Currently Syngenta's approach to the screening cascade, specifically deciding which compounds will pass each screen, is somewhat unscientific. It relies on a mixture of tried and trusted formulae and the ad hoc influence of the expert chemists and biologists, who might, for example, favour a particular compound and have reason for wanting it to pass further through the trial even though it performed poorly on an early test.

The problem is to develop methodology to optimize the screening cascade in a principled way, drawing on areas of Mathematics such as Statistics and Optimization.

1.3   Summary of our our approaches

During the study group week the team discussed the problem widely and considered a range of mathematical models for the screening cascade as well as different methods for choosing which compounds to pass through each screen.

EXT_ID	TEST#	Rate	ZEAMX	TRZAW	ORYSA	ALOMY	AVEFA	ECHCG	LOLPE	SETFA	ABUTH	AMAPA	AMARE	EPHHL	IPOHE	SOLNI	ST_ZEAMX	ST_TRZAW	ST_ORYSA	ST_ALOMY	ST_AVEFA	ST_DIGSA	ST_ECHCG	ST_ELEIN	ST_LOLPE	ST_SETFA	ST_ABUTH	ST_AMAPA	ST_AMARE	ST_CHEAL	ST_EPHHL	ST_IPOHE	ST_SOLNI
COMM4	TEST199	8						0	20	20		70		40	40							60	30	40	10	20		50		80	70	70	
COMM4	TEST200	8						0	0	0		70		40	50							70	40	40	10	10		40		80	80	80	
COMM4	TEST223	8						30	20	10			100		100							80			10	40			80			80	80
COMM4	TEST219	15						70	0	0			100		60	90						80			0	0			80			80	100
COMM4	TEST149	15.63	0			10		0		30	80		70								20		60				60	100		70			
COMM4	TEST216	15.63						40	0	10			100		90	100						70			0	20			90			90	90
COMM4	TEST218	15.63						70	0	0			100		60	100						70			20	0			80			80	90
COMM4	TEST199	30						80	10	0	100		100	90								70	70	90	0	30		80		80	60	70	
COMM4	TEST200	30						90	0	0	100		90	90								90	90	80	20	0		80			80	90	90
COMM4	TEST258	31						80	20	0		40		70	90								40		20	10		80				80	100
COMM4	TEST252	31.2	0					10		0	90	30		70		10							80			40	100	80			90		
COMM4	TEST225	31.25	0					70		0	100	100		90		10							80			40	80	80			90		
COMM4	TEST226	31.25	0					80		30	100	100		80		10							90			60	90	90			90		
COMM4	TEST227	31.25	0					70		30	90	100		100		0							90			50	90	80			90		
COMM4	TEST228	31.25	0					60		20	90	80		90		10							90			40	90	80			90		
COMM4	TEST229	31.25	0					10		10	90	70		80		0							80			40	90	80			80		
COMM4	TEST230	31.25	0					60		0	100	100		90		10							90			40	80	80			90		
COMM4	TEST231	31.25	10					70		10	100	100		80		10							90			30	90	80			90		

Figure 2: Snap shot of data recorded during screening process.

One aspect that struck all the mathematicians was the huge complexity of the process. The 'goodness' of a compound is difficult to characterize with a single variable and experiments on different compounds at the same level of screen are often conducted in a way that makes it difficult to compare their results directly, for example by their application at different concentrations. Figure 2 is snap shot of one of the tables of data collected during the screen: note in particular the large number of missing values.

In order to have somewhere to start with a preliminary analysis we chose to work with one of three highly simplified models of the screening process, although, all of which could be extended to be more realistic in a number of ways. The approaches taken in the four remaining sections are outlined briefly below.

In Sections 2 and 3 we introduce and analyze two models for the screening cascade that allow us to optimize the rate at which compounds are screened out at each stage. Both models use a simplified version of the problem in which there is a single 'goodness' score. The approach in section 2 uses Monte Carlo simulation and would be quite easy to adapt to include more complex modeling. The approach used in section 3 uses more probabilistic analysis and quickly computes very accurate results, but may be a little less general in its current form.

In section 4 we discuss the possibility of using prediction markets to incorporate the expert opinions of the chemists and biologists into the analysis in a mathematical way.

In Section 5 we introduce an even simpler model of the screening cascade, with only two levels of screen and binary experiment outcomes. This very basic model enables us to compute an optimal adaptive screening strategy, which makes optimal decisions about what tests to apply. For example choosing to terminate the whole process early if it will probably not be profitable.

2 Probabilistic model and Monte Carlo analysis

The problem outlined above can be characterized by the following simplified model set-up. We consider one shot of an annual cycle in which there are $j = 1, 2, \ldots, N$ compounds to evaluate in total: here $N = 1,000$. We suppose that each compound has a true (unknown) scalar merit μ_j, which rolls up a number of distinct attributes, such as lethality to various weed families and safety (i.e., non-lethality) to various crops. Of course, our approach might be extended in future to consider vector-valued merit and multi-criteria decision making.

The overall goal of the cycle is to try to identify the 'best' compounds, i.e., those with the highest merits, to carry forward to exhaustive field testing, which will measure compounds' merits rather accurately — in total 20 such compounds per year might be so selected.

The testing process is arranged into a sequence of $i = 1, 2, 3, 4$ stages, of increasing accuracy and expense, where $i = 4$ is the field test. The idea is that each stage is a screen which selects compounds to be passed to the next stage. In the real-world situation, compounds may skip stages if they appear especially promising, or may be re-introduced following rejection if other criteria (such as success of compounds with similar chemistry) indicate that they should. However, we neglect these difficulties for the time being. Let n_i be the number of compounds tested in stage i. We have

$$N(= 1000) \geq n_1 \geq n_2 \geq n_3 \geq n_4 \ (= 20). \tag{1}$$

If $n_i = n_{i+1}$ for any i, then stage i has not screened out any compounds at all — and thus is redundant.

In essence, the tests at each stage estimate the compounds' merits, but are erroneous because of either

- the simplified nature of the test (i.e., in the greenhouse, with only a limited number of weed and crop varieties) does not reflect the true requirements (in the field, with a wide range of uncontrolled environmental factors); or

- sampling error (i.e., there are not enough plants of each variety to draw robust statistical inference) — this is also convolved with measurement error due to the subjective human-based assessment used in each test.

The idea is to simplify these errors and suppose that each test returns the true merit plus a random variable, for example we use

$$\mu_{ij} = \mu_j + N(0, \sigma_i^2), \tag{2}$$

where μ_{ij} is the estimate of compound j's merit at stage i and $N(0, \sigma_i^2)$ represents a normally distributed random variable whose standard deviation $\sigma_i \geq 0$ is a measure of the
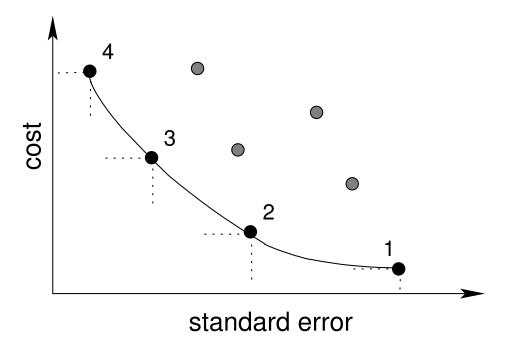
Figure 3: Trade-off between the cost (c) and the standard error (σ) of a test. Tests that are used in practice (black discs), labelled 1,2,3,4 to represent their stage in the screening process, should lie on a Pareto front amongst the set of all possible tests. Inferior tests (grey discs) lie behind the Pareto front and should not be used.

accuracy of testing in stage i. Smaller values of σ_i represent more accurate tests and the limit $\sigma_i = 0$ represents a perfect test that recovers the true merit. Note we assume that each stage is an unbiased estimator (so we assume that a calibration has been carried out) and that there is no correlation between errors from stage to stage. Note

$$\sigma_1 > \sigma_2 > \sigma_3 > \sigma_4, \tag{3}$$

expresses the increasing accuracy of tests from stage to stage.

We let c_i be the cost of testing a single compound at stage i, and we assume that this increases from stage to stage so that $c_1 < c_2 < c_3 < c_4$, in contrast to (3). In fact, it is instructive to understand the (σ, c) plane, see Fig. 3. In particular, the tests (σ_i, c_i) used in practice should lie on a Pareto front in the set of all possible tests. It is also instructive to consider a setting in which sampling error is the only kind. In this, the cost of each test is proportional to its size (the number of plants used), so the central limit theorem will imply that c is proportional to $1/\sigma^2$.

Depending on the precise objective, there are several different problem formulations, but here we suppose that there is a fixed budget $B > 0$ which must be fully spent. We also suppose that n_4 (the number of compounds taken to field test) is fixed, absorbing $n_4 c_4$ of the

budget. The decision variables are thus the proportions $0 \leq p_i \leq 1$ of the remaining budget $B - n_4 c_4$ spent on stages $i = 1, 2, 3$. Thus

$$p_i = \frac{c_i n_i}{B - c_4 n_4}. \tag{4}$$

and

$$p_1 + p_2 + p_3 = 1. \tag{5}$$

The latter equation implies there are only two degrees of freedom to investigate in a design space which is best viewed in Barycentric coordinates, see Figs. 4 and 5.

In order to compare the efficacy of different screening systems, we use an objective function which measures the relative ranking of the compounds that are passed to stage 4. Let R be a bijective map on $\{1, 2, \ldots, N\}$ that describes compounds' rankings, based on true merit, so that $\mu_{R^{-1}(1)} \geq \mu_{R^{-1}(2)} \geq \ldots \geq \mu_{R^{-1}(N)}$. Let $i_1, i_2, \ldots, i_M$ be the compounds that pass to stage 4 (where $M = n_4 = 20$). Then we define

$$f = 1 - \frac{\frac{1}{M} \sum_{j=1}^{M} R(i_j) - \frac{M+1}{2}}{N - M - \frac{1}{2}}. \tag{6}$$

We thus derive a perfect score of $f = 1$ if the best 20 compounds, based on true merit, are selected. In contrast, $f = 1/2$ corresponds to a random selection and $f = 0$ to when the worst 20 compounds are selected.

Our procedure is then to test each screening design, represented by the vector $\mathbf{p} = (p_1, p_2, p_3)$ satisfying (4,5), by Monte Carlo methods. To do this, we create a collection of N artificial compounds whose true merits are drawn from a probability distribution and are hidden from the screening process that follows. Each stage's tests are then simulated in turn by adding the appropriate noise, as defined above, to the true merit. The compounds with the best simulated merits are then passed to the next stage of screening.

The overall quality of the cascade is then measured according to the rank function f of the compounds that pass to stage 4. Note that f is a random variable because of the noise incorporated in each test. Hence to obtain stable expected values, the whole procedure is ensembled a large number of times and the results averaged. Of course, one might attempt an analytical approach (see Section 3), but the Monte Carlo method runs extremely quickly — typically just a few seconds to explore the design space and produce each of Figs. 4 and 5 — and can be programmed extremely flexibly to adapt to more sophisticated rules (e.g., compounds that skip stages).

There is generally speaking a huge variety in the possible solution structures, depending how the parameters σ_i, c_i ($i = 1, 2, 3$) and B are chosen. Furthermore, the standard deviation σ of the true merit distribution is a key parameter, and if it is comparable or less than the standard error of the tests, then the performance f will be extremely poor (although one

might argue that the ranking metric is not relevant if compounds' true merits are very similar).

Figs. 4 and 5 present exemplar results for high and low budget scenarios respectively. In each case, the parameters σ_i, c_i and σ are set at the same 'stylised' values (not formally fitted to experimental data), and we report the average f from 100 simulations for each feasible data point $\mathbf{p}$. Note large parts of the barycentric triangle are blank, because they correspond to settings which do not satisfy (1), which requires that the number of compounds tested is reduced from stage to stage.

In the high budget scenario, we set $B = 2B_{\min}$, where $B_{\min} = n_4 \sum c_i$ is the total cost of carrying 20 compounds through all of the stages. Fig. 4 shows that in this case, it is best to test all compounds in the first stage ($n_1 = N$), and the search for the optimal strategy is thus a line search that explores the trade off between the cost of the second and third stages. Of course, the optimal trade off may be read off from the simulation output (indicated by a black circle on the plot).

In the low budget scenario, we set $B = 1.05B_{\min}$, so that there is almost no spare resource. In particular, it is only just about feasible at our parameter settings to test all compounds in the first stage, and this would leave almost no resource for later stages. It follows that the best thing to do is to randomly select a subset of compounds to test in the first stage and thus we need to explore three-way trade offs between the costs of the stages 1, 2, and 3. (From the wider system design point of view, this finding has implications for how the testing budget is balanced with the budget allocated to compound discovery.) As before, the optimal trade off may be read off from the simulation output (indicated by a black circle on the plot). However, note, as one might expect, the performance of the cascade, as indicated by the maximum value of f, has been severely impacted by the constrained budget, in comparison to Fig. 4.

Finally, note that both sets of results indicate very poor performance along the lines $n_1 = n_2$ (resp. $n_2 = n_3$). This is because these describe settings in which there is no cut-down from the first stage to the second stage (resp. second stage to the third stage) and thus the first stage (resp. the second stage) is redundant and the resource spent on it has been wasted. This finding suggests exploring set-ups in which the first stage (resp. second stage) are deleted. This kind of question, relating to the number of and structural relationships between the stages, remains for future work.
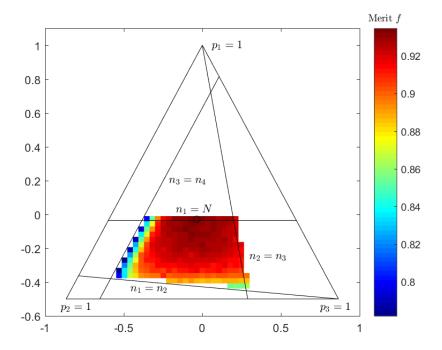
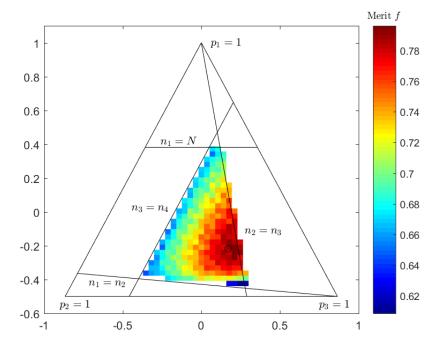Figure 4: Monte-Carlo results for a high budget scenario.

Figure 5: Monte-Carlo results for a low budget scenario.

3 Probabilistic approach (analytical)

Each of the experiments in the screening cascade is designed to test how effective is the molecule going to be in the final application. The experimental results however always consist of a random error. One can try to simulate the screening cascade using Monte Carlo methods, by randomly generating each time the score reached by each of the molecules. These methods however require to repeat this procedure multiple times in order to achieve a reliable results.

Another approach is to compute the probability distribution change during each process using basic mathematical transformations. Its biggest advantage is no need to repeat computations, very high precision achievable even with very low computational power usage as well as high possibilities of manipulating the screening procedure. This approach is presented in the section 3.1, while in section 3.2 the results of numerical experiments based on this model are presented. Section 3.3 shows how can the model be used to verify how could data aggregation across the experiment improve the overall screening cascade performance.

3.1 Probabilistic model of the screening cascade

There are three assumptions of the model:

1. Initial frequency distribution of the true merit is known; it will be referred as $\mathbb{P}_{\text{initial}}(x)$.

2. Probability distribution of experimental results around real value is known and is constant for the tested population; it will be referred to as $\mathbb{P}_{\text{experiment}}(x)$.

3. Results of all experiments are independent from each other, i.e. the random error of each experiment is not correlated with those of others experiments.

The term *frequency distribution* is used for either a discrete or continuous distribution, which sum or integral in the given range is interpreted as the expected number of samples (molecules) with values included in this range. The main point of modeling the screening process is to check how the initial frequency distribution change during a single experiment. The procedure for finding the frequency distribution after passing through a screen is composed of three steps:

Step 1. Find the frequency distribution of the measured values by calculating convolution of

the distributions $\mathbb{P}_{\text{initial}}$ and $\mathbb{P}_{\text{experiment}}$, i.e.:

$$\mathbb{P}_{\text{measured}}(x) = (\mathbb{P}_{\text{initial}} * \mathbb{P}_{\text{experiment}})(x)$$
$$= \int_{-\infty}^{+\infty} \mathbb{P}_{\text{initial}}(y)\mathbb{P}_{\text{experiment}}(y - x)\mathrm{d}y. \tag{7}$$

Step 2. Find a threshold value, such that measurements of exactly m samples are placed below this value. These samples are going to be removed during the screening process.

$$\mathbb{P}_{\text{measured}}(X <= \text{threshold}) = m. \tag{8}$$

Step 3. Find the final distribution of true merit across the population, which has advanced for the next step of screening cascade. It can be found by multiplying the initial frequency distribution by the probability that the measured value is higher or equal to the threshold:

$$\mathbb{P}_{\text{final}}(x) = \mathbb{P}_{\text{initial}}(x)\mathbb{P}_{\text{experiment}}(X >= \text{threshold} - x). \tag{9}$$

The distributions found during an example of the following procedure are presented on the Figure 3.1.
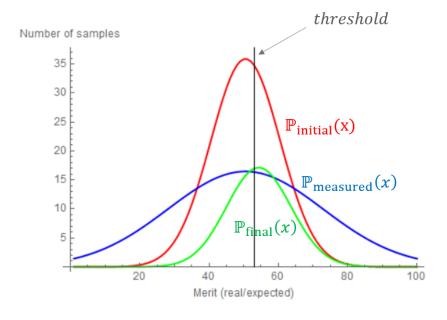

Figure 6: Distributions in case of $\mathbb{P}_{\text{experiment}} = N(0, 20)$ and $\mathbb{P}_{\text{initial}} = 900 \cdot N(50, 10)$

To model the whole screening cascade one needs to repeat the presented procedure for each experiment, in each step choosing the final distribution from the previous step as an initial population of the next step. As the aim of any optimization procedure one can try

to maximize the expected value of true merit among the sample, which reached the field tests stage. Mathematically it is expressed as:

$$\mathrm{E}(X) = \frac{\int_{-\infty}^{+\infty} x \mathbb{P}_{\mathrm{final}}(x) dx}{\int_{-\infty}^{+\infty} \mathbb{P}_{\mathrm{final}}(x) dx}. \tag{10}$$

The equations presented above cannot be solved analytically and there is often no simple mathematical expression for the frequency distribution $\mathbb{P}_{\mathrm{final}}$. This model can however be implemented by splitting a given range of merit values into a sections and by assigning to each section the probability that a sample has merit within this range. As long as the probability distribution decreases exponentially for low and high merit values (i.e. it doesn't have so called heavy tail) one can limit the domain from $(-\infty, +\infty)$ to some finite range where we expect to have merit values. However one needs to remember to keep the distributions properly normalized.

3.2   Numerical experiments based on the model

The described model can be used to design the experimental procedure, i.e. which experiments are going to be used, in which order and how many sample should be dropped during each step of the screening cascade etc. One can introduce additional constraints, for example fixing the total cost of all experiments. Some proposition of numerical experiments are listed below:

1. Varying the number of trials in each screening process in order to find the best configuration at a given budget,

2. Adding, removing or manipulating the order of the tests in order to find their impact on the final results,

3. Finding the optimal number of initial samples, which are tested, in order to check whether taking all possible molecules for initial tests is optimal.

The example results of the first proposition are presented in this subsection. Firstly, some assumptions about $\mathbb{P}_{\mathrm{initial}}(x)$ and $\mathbb{P}_{\mathrm{experiment}}(x)$ for each of screening cascade stages has to be made. Here it is assumed that all the distributions can be modeled as normal distribution $N(\mu, \sigma)$. Moreover it is assumed that exactly 20 of the best samples from the last glasshouse experiment are taken to the field tests.

The initial distribution of true merit is modeled as $\mathbb{P}_{\mathrm{initial}}(x) = 900 \cdot N(50, 10)$, where 900 is the initial population size. Note that the choice of expected value does not influence the

order of samples, which is the only important feature during selection. It is chosen in order to have any absolute scale for its value. It does not even need to be used in the data presentation, because we can present merit values in comparison to the mean of the 20 best samples from the initial population (which can be regarded as the upper limit for any testing algorithm). This mean can be found by calculating the following expression:

$$\mu_{\text{max}} = \frac{1}{20} \int_{\text{threshold}}^{\infty} x \mathbb{P}_{\text{initial}}(x) \mathrm{d}x, \tag{11}$$

where threshold can be found by solving the following equation:

$$\mathbb{P}_{\text{initial}}(X > \text{threshold}) = 20. \tag{12}$$

The standard deviation of distributions is however important, because it allows us to compare the precision of each test to the diversity of the initial population. Each experiment is modeled using the distributions presented in Table 3.2. Zero expected value means that the experiment is unbiased. The cost is not needed to apply the presented model, but it allows us to draw conclusions about the best design of experiment under budget constraint.

experiment	distribution	cost per sample	number of removed samples
EPS	$N(0, 20)$	1	m_1
PPS	$N(0, 4)$	10	m_2
SPS	$N(0, 2)$	6	$880 - m_1 - m_2$

Table 1: Properties for each screening process used in numeric experiment.

From now on, the impact of the number of samples dropped in each stage (m_1 and m_2) on the quality of the final set chosen for a field tests is investigated. The developed model was applied for $m_1 \in 0, 10, 20, 30, ..., 880$ and $m_2 \in 0, 10, 20, 30, ..., 880 - m_1$ in order to find the mean value of true merit among 20 samples chosen for the field tests. Figure 3.2 shows how the frequency distribution changes in each stage of screening cascade for $m_1 = 500$ and $m_2 = 280$. As could be predicted, the mean true merit increases in each stage, however also the number of samples with the highest merit values slightly decreases.

Figure 3.2 presents the results both for the mean true merit value and the total cost of the experiments. The value of true merit is expressed as $\log_{10} \mu_{\text{max}} - \mu(X)$, so that the lower values are closer to the optimal solution μ_{max}.

It can be easily read from the graph that the best solution is to leave all the samples until the most precise SPS test ($m_1 = 0$ and $m_2 = 0$), however this procedure is obviously the most expensive one. The cheapest procedure is to drop almost all the samples in the first test ($m_1 = 850$), but it is obviously the least precise method. To obtain more useful conclusions
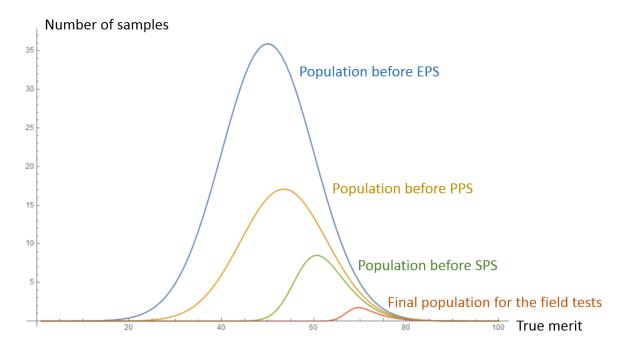
Figure 7: Frequency distribution of population left after each stage of the screening cascade.
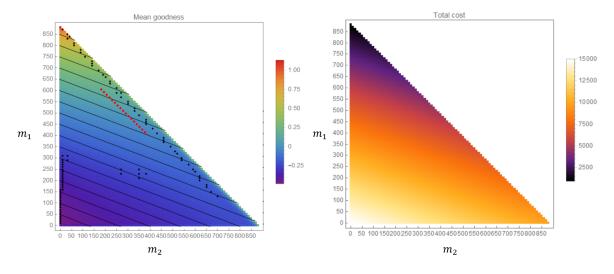

Figure 8: Value of $\log_{10} \mu_{\max} - \mu(X)$ (left) and total cost (right) for different procedure settings.

the lines of the fixed income, described by the linear equation $1 * 900 + 10 * (900 - m_1) + 6 * (900 - m_1 - m_2) = \text{const.}$ were added. For each budget the optimal configuration was found and is represented by the black points.

What can be noticed is that for high budget it's good not to remove any or almost any samples in the second test (PPS). Basically, as long as the assumptions are reliable this step could be removed from the procedure. However the situation changes dramatically when the budget is below the certain threshold, where a phase transition takes place. For lower budget (which is the area of our interest) both EPS and PPS are used to reduce the number of samples to very small quantities ($m_1 + m_2 \in [800, 880]$). It means that only a few samples should advance to the last test (SPS), however according to the model this step should not be removed from the procedure. It provides an extra validation step, while taking up only a few samples just slightly increasing the costs.

The red line on Figure 3.2 corresponds to the selection currently used by the Syngenta, so if the assumptions about distributions are reliable, it is suggested to slightly limit the number of samples taken to SPS in the same time increasing the number of samples passing the less reliable EPS. The graph in Figure 3.2 present the outcome of the optimal procedure for different values of the available budget. Such a graph can be used as an additional support in making decisions about increasing and decreasing the budget for the screening cascade.
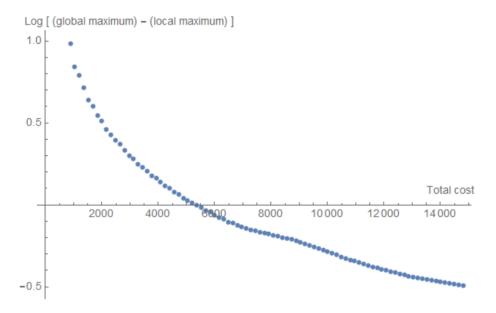

Figure 9: The dependence between the total cost and the mean true merit of final population in case of optimal configuration.

This completes an example of using the presented model. Before applying a similar procedure, it is recommended to gather some data about the variance of initial population and each of the experiments. However it is difficult to do unless some randomly chosen subset of molecules is tested at each stage of the screening cascade and compared to their results at the field tests. Another option is to use heuristics to propose some arbitrary

distributions as it was done while working on the presented example.

3.3 Aggregating the experimental data

One of the questions the study group was asked was to consider aggregation of data from all experiment during selection of samples passing to the next stage of the screening cascade. For example if a molecule has passed EPS and PPS with a very high score, but obtained a moderate score in SPS (not included in the top 20 samples) it may still be worth taking to the field experiment. The solve this problem two questions need to be addressed:

- How should the result for different results be aggregated?

- How would it increase the quality of the samples taken to the field test?

Using only the assumptions presented in section 3.1 both of these questions can be answered. If the results of the experiments are independent from each other (i.e. there is no correlation among them) the best estimate of the mean is expressed as:

$$\mathrm{E}(X) = \frac{\sum_i \frac{x_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}}, \tag{13}$$

where i is the index of experiments, which were already conduced (e.g. EPS is 1, PPS is 2, SPS is 3), x_i is the result of ith experiment and σ_i is the standard deviation of $\mathbb{P}_{\text{experiment}}$ distribution for i^{th} experiment. The variance of this estimator is:

$$\mathrm{Var}(X) = \frac{1}{\sum_i \frac{1}{\sigma_i^2}}, \tag{14}$$

which, as long as more than one test was conducted, is always smaller than the variance of any single experiment σ_i. This improves the precision of each screening stage. For the distributions presented in Table 3.2, one can find the effective precision of each stage, which is presented in the Table 3.3. The more test are already conducted the precision is further improved.

	EPS	PPS	SPS
Original standard deviation	20	4	2
Aggregated standard deviation	20	3.92	1.78

Table 2: Precision comparison for aggregated and not aggregated score in each stage.

The experiment described in subsection 3.2 were repeated again for improved precision. Figure 3.3 presents the improvement of true merit mean among 20 samples taken to the final tests. It's not clear whether the waves seen on the graph are reliable or whether they are just caused by the numerical imprecision, however some general trends can be observed. The improvement is the highest when a large set of samples reach the SPS test, which happens for high budget strategies. In low budget case, when only a very small number of samples reaches the SPS test the improvement is low.
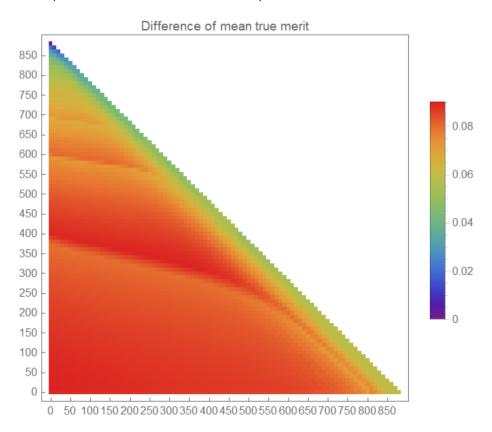

Figure 10: The comparison of screening wit and without aggregated for different budgets.

This is even more clear if the precision vs cost curve is drawn both for the results with and without aggregation as presented on the Figure 3.3. It is clearly seen that for the low budget the improvement is small, however it has to be emphasized that the aggregation of data does not require any modifications of the experimental procedure (such as changing the number of investigated samples). In fact, if the assumptions are reliable the effect of aggregating data is almost the same as the effect of optimizing the procedure as presented in the previous section. The only disadvantage of this modification is however a need to estimate the precision of each test accurately in order to implement the formula for score aggregation (equation 13). Choosing unreliable variances could lead to much lower improvement or even to deterioration if fist stages would get too high weights.
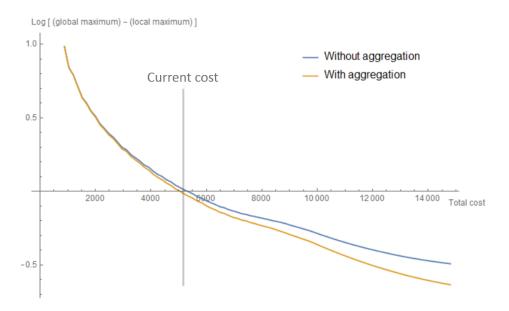
Figure 11: The comparison of screening wit and without aggregated for different budgets.

4   Prediction markets for cascade screening

While working on the problem different mathematical approaches for the modelling were presented. However there are some problems, which cannot be resolved using standard mathematical tools. These are:

- heuristics involved in the process, especially in the early stages of the molecules development,

- different perspective and objectives of people involved in the process (chemists, biologists and others),

- the problem of insufficient data, while increasing the scale of the process would be too costly.

A possible solution to address all these problems is applying crowdsourcing using prediction markets.

4.1   Description of prediction markets

We will start with the definition given by Leigh and Wolfers: prediction markets are markets where participants trade contracts whose payoff depends on unknown future events. The defining feature of a prediction market is that the price of these contracts can be directly

interpreted as a market-generated forecast of some unknown quantity [1]. Their mechanism relies on the efficient markets hypothesis: the price of a financial security or prediction market contract reflects all available information [1]. Therefore prediction markets are an example of efficient crowdsourcing - aggregating dispersed, and often contradictory, knowledge from a group of people to obtain very precise information about the outcome of a future event.

Modern approach to prediction markets began in 1988, when three economists of the Iowa University created a market to predict the outcome of the presidential election (Bush vs Dukakis). It was observed that in any given moment in time such market gave much better forecast than all major polls. The experiment has been carried on for many other elections and the comparison shows that it beats all polls in about 75% of the times. The advantage of markets is even bigger when the time to election is long. [2]

After the success of Iowa Electronic Markets the interest in prediction markets grew rapidly. Currently markets are used by many large corporations (Google, Microsoft, IBM, Lockheed Martin, etc.) as a tool to assess the probability that a project will end as planned, that a sales goal will be achieved or as a tool to estimate the market potential of innovative products [3]. For instance, General Electric has been running markets for new ideas and products originated by employees. Eli Lilly, a large pharmaceutical company, ran a prediction market to support choice of new drugs for further development, primary decision factor being market potential. Further, in the BRAIN[1] project - an internal research at Hewlett-Packard, it was shown how to run prediction markets with small numbers of participants (up to 10 people) and still obtain meaningful results.

Apart from internal corporate applications there are also many publicly available commercial markets (e.g. Hollywood Stock Exchange, Intrade), where operators often profit from fees or selling complex analyses derived from the market data. Furthermore, even DARPA[2] and IARPA[3] have implemented prediction markets [4], mainly to obtain accurate predictions important for the American military or intelligence community.

4.2 Enhancing information for cascade sceening

Prediction markets are primarily a source of information that can be aggregated efficiently from their participants. This property allows to use them to reduce the uncertainties in the process. For instance, the information that could possibly be gathered through a prediction market encompass the progression of samples through the screening cascade, the

---

[1]   Behaviorally Robust Aggregation of Information in Networks

[2]   Defence Advanced Research Projects Agency - an American government agency supporting large scientific projects that might be useful for military purposes

[3]   Intelligence Advanced Research Projects Agency - the counterpart of DARPA devoted to intelligence purposes

possible paths of molecules development and success of new projects can be predicted.

Another very important benefit from the use of prediction markets is the fact that all parties, including chemists, biologists and others would be simultaneously involved in the planning process. While interacting on one platform everyone could feel that his voice is heard and that he can have impact on the entire process.

For instance, take one of the biggest uncertainties in the planning process - the selection of the samples, which are tested using the screening cascade. As we learned, biologists and chemists often have different perspectives on which properties of the molecules has highest impact on their success during glasshouse and field tests. This difference may lead to different preferences for the molecules development. For example there are some additional factors best known to the biologists observing the effect of pesticides on the plants, which are not considered by the chemists preparing these molecules in the laboratory. All this data is of great significance for predicting the success of the samples throughout the screening process.

All this information can be easily aggregated into forecasts by a properly set prediction market among people involved in the process. Apart from chemists and biologists working directly on the experiments, others participants are also welcome, as they increase diversification of information that in turn can enhance accuracy of a prediction market ([1] and [3]). The questions on the market, that have to be binary, could ask about the success of a molecule or group of molecules in each step of the screening cascade separately. As an illustrative example we provide a set of questions, which could be put on a prediction market:

1. Is ⟨given sample⟩ going to reach and pass the PPS test?

2. Are at least five sample containing ⟨given compound⟩ going to pass through PPS test?

3. Are there going to be more sample containing ⟨given compound⟩ than the ⟨other compound⟩ among those, which reach the PPS test?

Similar sets of questions could be posed for every stage of the screening cascade or any samples and molecules. Obtained predictions may turn out to be more accurate than the early stages of molecule production and testing process, especially those based on the heuristics of people working on their production.

Such a prediction market could be incorporated into a larger software tool that would help in the screening cascade planning process and would allow for fast and efficient information exchange between all parties involved. On the other hand data obtained on such market would allow to build complementary datasets, involving information about the predicted success of passing expensive tests (for example field tests) even if these molecules

are not going to reach this stage. It can be used to calibrate other mathematical models described in this report, for example by estimating the accuracy of early stages of the screening cascade.

5   Optimal Adaptive Screening

In practice the success rates associated with each screen and the information that one screen gives about the next are unknown. Instead they are best thought of as random variables and reasoned about using Bayesian statistical methods. We might believe that a good score in test 1 will result in a good score in test 2 but unless there has been a rigorous study designed specifically to explore this relationship we will not be certain about what probabilities to use when choosing our screening cascade. Furthermore, even if these probabilities are well known for one group of compounds, they may not be applicable to a new group developed from a different chemical lead.

Instead we propose an adaptive screening strategy, which chooses compounds for testing at different levels of screen and continually updates its strategy based on the results that it has seen so far. If we assign a financial cost to each level of screen and a financial reward to finding compounds that perform well in the final field trials then the whole screening process can be formulated as an optimization problem. In particular we want to maximize the expected profit subject to various possible constraints.

To do this we model the whole screening process as a Markov Decision Process (MDP). A MDP consists of a state space X an action space A and a reward function R. In the case of the screening process the state space X consists of every possible configuration of tests performed and their results. For example a state $x \in X$ could represent that we have carried out test 1 on compound 1 and it scored 7, we carried out test 1 and test 2 on compound 2 and it scored 4 and 6 respectively, and we didn't carry out any other tests. The action space represents of all of the possible actions that we can take from each state. For example from the state x we could carry out test 2 on compound 1, or we could carry out test 3 on all of the compounds etc... The rewards represent the profit associated with each state. This should be equal to the reward associated with any compounds that we have found to pass the final test minus the cost of all of the tests that we have carried out.

The MDP model is Bayesian in the sense that given that we are in a state $x \in X$, we have some information about the underlying probabilities. In particular we have the test scores for all of the tests that we have carried out so far. For example if we have carried out test 1 a number of times then we will have data about the test 1 performance scores, which will reduce our uncertainty about the probability distribution of the test 1 performance scores. This knowledge gained through the process is exploited by the algorithm to find the globally optimal adaptive screening strategy.

The only drawback to the MDP formulation is that the size of the state space grows exponentially with the number of compounds and screens. This means that we are not practically able to use this formulation to optimize the screening strategy for a model that is remotely

faithful to the real life screening problem that Syngenta work with. However, there are a variety of extensions and alternative to MDP, such as policy iteration with state-action value function approximation, which are often used to overcome this problem. In applications of these other techniques an important first step is always to study an MDP formulation of a simpler and smaller version of the original problem. Exploring the MDP model of the toy problem lets us gain an understanding of the different factors at play in determining the optimal adaptive screening strategy, which can then be incorporated into a more complex model. Such a preliminary study is the focus of the remainder of this section.

5.1 Binary two level screen

We consider a screening problem consisting of $n = 20$ candidate compounds and two levels of testing. Each test has a binary outcome 0 meaning a fail or 1 meaning a pass. Thus each compound can have one of four possible performance profiles $(0,0)$ meaning that it fails both tests, $(0,1)$ meaning that it fails test 1 and passes test 2, $(1,0)$ meaning that if passes test 1 but fails test 2 and $(1,1)$ meaning that it passes both tests.

We assume that there is an underlying probability distribution $p = [p_{00}, p_{01}, p_{10}, p_{11}]$, such that p_{00} is the probability that a randomly chosen compound will have performance profile $(0,0)$ and so on. However, rather than assuming that we know p exactly we instead suppose that it is an unknown random variable but with a known prior distribution. In particular we will assume that p has a Dirichlet distribution with concentration parameters $[a_{00}+1, a_{01}+1, a_{10}+1, a_{11}+1]$. The choice of a will bias our prior in different directions. One way to interpret the concentration parameters is to imagine that we began with a uniform prior, which assumes that any p is equally likely - including even p with test 1 and test 2 negatively corrolated, and then tested a total of $a_{00}+a_{01}+a_{10}+a_{11}$ randomly chosen compounds and found a_{00} of them to have performance profile $(0,0)$ and so on. Given the choice for a the average probability of a randomly chosen compound passing test 2 is equal to

$$\lambda = \frac{a_{01}+1+a_{11}+1}{a_{00}+1+a_{01}+1+a_{10}+1+a_{11}+1},$$

and the average probability of a randomly chosen compound passing test 2 given that it passed test 1 is equal to

$$\mu = \frac{a_{11}+1}{a_{10}+1+a_{11}+1}.$$

However, because the probabilities are themselves random variables, the screening strategy will need to be able to deal with variations around these averages, which it will not be aware of until it has carried out some screens and gathered some data.

From the screening process's perspective each compound can be in one of nine possible states $(?, ?)$ meaning that no tests have been carried out, $(0, ?)$ meaning that we have carried

out test 1 and it failed but we have not carried out test 2, and so on. The full state space of the screening process can therefore be represented by

$$X = \{(?,?), (0,?), (1,?), (?,0), (0,0), (1,0), (?,1), (0,1), (1,1)\}^{20},$$

which assigns each of the 20 compounds to one of the nine possible states. The number of different states in X has twenty digits! Luckily we can reduce this considerably without introducing any approximation. Since each compound is a *priori* i.i.d. all that matters is the number of compounds we have in each of the nine possible states. Thus the state space can be represented by $X = \{x \in \mathbb{N}^9 \; : \; \sum_{i=1}^{9} x_i = 20\}$. The state $x \in X$ represents having x_1 compounds in state $(?,?)$, x_2 compounds in state $(0,?)$ and so on. By representing the state space in this way we reduce the number of states to a large but manageable $3,108,105$.

For each state there are at most six possible actions. If the state contains at least one $(?,?)$ compound then we can apply test 1 or test 2 to it. If the state contains at least one $(?,0)$ compound then we can apply test 1 to it, if the state contains at least one $(?,1)$ compound then we can apply test 1 to it, if the state contains at least one $(0,?)$ compound then we can apply test 2 to it and if the state contains at least one $(1,?)$ compound then we can apply test 2 to it.

We compute the optimal adaptive screening strategy by solving the Bellman equations using dynamic programming. We compare the optimal adaptive screening strategy obtained with the MDP formulation to two simpler non-adaptive methods. In non-adaptive method 1 we simply skip test 1 and apply test 2 to every compound. In non-adaptive method 2 we first apply test 1 to every compound and then apply test 2 to every compound that passed test 1.

5.2 Experimental Results

We will assume that the cost of test 1 is £0.2 and the cost of test 2 is £1. For this example we use the concentration parameters with $a = [10, 1, 1, 1]$ meaning that the average probability of a randomly chosen compound passing test 2 is $\lambda = 4/17$ and the average probability of a randomly chosen compound that passes test 1 passing test 2 is $\mu = 1/2$. We will vary the reward for finding a compound which passes test 2 between £2 and £6. The results are displayed in Figure 12. Note that with a low reward both non-adaptive methods expect to make a loss. The optimal adaptive method will never have a negative expected profit, although it is still possible for it to make a loss in a single trial if it 'gets unlucky'. Over roughly £4.75 non-adaptive method 1 outperforms non-adaptive method 2. This is because the reward is so large that it is on average worthwhile screening all of the compounds in test 2. However the optimal adaptive method always makes a bigger profit.
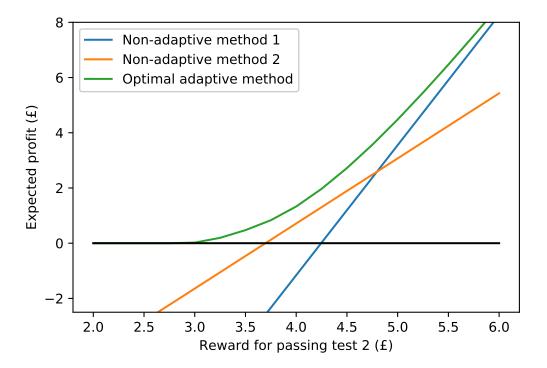
Figure 12: Comparison of screening strategies using Dirichlet prior with $a = [5, 1, 1, 2]$.

5.3   Optimal adaptive screening behavior

As well as being able to compute summary statistics for the performance of the optimal adaptive screening method we can simulate randomly generated trials and observe the optimal decision making process in action. In further work we could represent the method's behavior as a decision tree and compute the probabilities associated with each branch. But for this report we simply generate a few random trials. For all the trials below we use the strategy that is optimized for a reward of $£4$ for each compound that passes test 2. For each example the exact sequence of states is depicted in Section 7.

Interestingly the optimal adaptive screen always starts by testing the first compound with test 2. If this compound fails as in Examples 1 and 2 then the method begins applying test 1 to the compounds.

In Example 1 the method finds only one compound that passes test 1 out of five and this compound fails test 2. Given the evidence that few compounds pass test 2 and that test 1 is a poor predictor of test 2, the method chooses to stop early.

In Example 2 the method also initially finds only one out of five compounds pass test 1 but this compound also passes test 2. Given the evidence that test 1 is a good predictor of test 2, the method then applies test 1 to every compound and then applies test 2 to every compound that passed test 1.

In Examples 3, 4 and 5 the first compound chosen passes test 2. In these examples the method then proceeds to apply test 2 to more compounds. In each of these examples the second compound fails test 2 and the method then applies test 1 to the compound that passed test 2.

In Example 3 the first compound fails test 1. The method then applies test 2 to more pre-viously untested compounds and only one out of six compounds pass. Given the evidence that few compounds pass test 2 and that test 1 is a poor predictor of test 2, the method chooses to stop early.

In Example 4 the first compound fails test 1. The method then applies test 2 to more previ-ously untested compounds and the finds that three out of six compounds pass. Given the evidence that many compounds pass test 2 but that test 1 is a poor predictor of test 2, the method continues to test all of the remaining compounds with test 2.

In Example 5 the first compound passes test 1. Given the evidence that test 1 is a good predictor of test 2, the method then applies test 1 to every compound and then applies test 2 to every compound that passed test 1.

5.4 Discussion

We have presented a MDP model for a simple two stage screen with binary outcomes. The MDP formulation enables us to compute the optimal adaptive screening strategy, which is designed to maximizes the expected profit. We hope that this preliminary study can form the basis of further work, developing methodology to compute optimal adaptive screening strategies for more complex screening models and ultimately to be applied to Syngenta's herbicide screening program. Some of the additional factors that would need to be include in such a model include.

- Much larger number of compounds: typically there are around 1000 compounds in each screen.

- Additional information about each compound: physical or chemical characteristics could enable more accurate predicted performance.

- More screening stages: the Syngenta screening process consists of around six stages.

- Continuous, possibly vector valued, screening performance scores: the scores of each screen are kill rates on various types of plants.

- Time and logistical constraints: it is not practical to carry out one test at a time. Instead the tests should be done in large batches, but not to exceed the capacity of the laboratory/greenhouse available.

In principal all of these considerations could be put into a MDP model, but the resulting compilation would be beyond the powers of current computers. However, developing such a MDP model and then exploring alternative approximate approaches to optimizing it would be a realistic approach to developing an optimal adaptive screening method for Syngenta's herbicide screening program.

6   List of Acronyms

References

[1] Leigh, Wolfers, Prediction markets for business and public policy, The Melbourne Review, 3 (1), May 2007.

[2] Stix, Super Tuesday: Markets predict outcome better than polls, Scientific American, 4 February 2008.

[3] Ho, Chen, New Product Blockbusters: The Magic and Science of Prediction Markets, California Management Review, 50 (1), 2007.

[4] http://www.goodjudgmentproject.com/, http://hanson.gmu.edu/policyanalysismarket.html

7 MDP Example behavior

7.0.1 Example 1

Matrix 1:

$$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 2:

$$\begin{bmatrix} ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 3:

$$\begin{bmatrix} ? & 0 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 4:

$$\begin{bmatrix} ? & 0 \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 5:

$$\begin{bmatrix} ? & 0 \\ 0 & ? \\ 1 & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 6:

$$\begin{bmatrix} ? & 0 \\ 0 & ? \\ 1 & 0 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 7:

$$\begin{bmatrix} ? & 0 \\ 0 & ? \\ 1 & 0 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

Matrix 8:

$$\begin{bmatrix} ? & 0 \\ 0 & ? \\ 1 & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}$$

7.0.2 Example 2

$$
\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
$$

$$
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
$$

$$
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
$$

$$
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ 1 & ? \end{bmatrix}
\begin{bmatrix} ? & 0 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 0 & ? \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ 1 & 1 \end{bmatrix}
$$

7.0.3 Example 3

$$
\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ ? & 0 \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
$$

7.0.4 Example 4

$$
\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 1 \\ ? & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 1 \\ ? & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 1 \\ ? & 1 \\ ? & 0 \\ ? & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 0 & 1 \\ ? & 0 \\ ? & 0 \\ ? & 1 \\ ? & 1 \\ ? & 0 \\ ? & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
$$

```
[ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ] [ 0 1 ]
[ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? ? ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? ? ] [ ? ? ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? 1 ] [ ? 1 ] [ ? 1 ] [ ? 1 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? 0 ] [ ? 0 ] [ ? 0 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? 0 ] [ ? 0 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? 1 ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ]
[ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ] [ ? ? ]
```

```
[ 0 1 ] [ 0 1 ]
[ ? 0 ] [ ? 0 ]
[ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ]
[ ? 1 ] [ ? 1 ]
[ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ]
[ ? 0 ] [ ? 0 ]
[ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ]
[ ? 1 ] [ ? 1 ]
[ ? 0 ] [ ? 0 ]
[ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ]
[ ? 1 ] [ ? 1 ]
[ ? 0 ] [ ? 0 ]
[ ? 0 ] [ ? 0 ]
[ ? 1 ] [ ? 1 ]
[ ? 1 ] [ ? 1 ]
[ ? ? ] [ ? 1 ]
```

7.0.5 Example 5

$$
\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 1 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} ? & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ 1 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ 1 & ? \\ 0 & ? \\ 0 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ 1 & ? \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
\begin{bmatrix} 1 & 1 \\ ? & 0 \\ 1 & 1 \\ 0 & ? \\ 0 & ? \\ 1 & ? \\ 1 & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}
$$

Top row of matrices (left to right):

```
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 0]  [1 0]  [1 0]  [1 0]
[1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 1]  [1 1]  [1 1]
[? ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[? ?]  [? ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]
[? ?]  [? ?]  [? ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [0 ?]  [0 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [0 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]
```

Bottom row of matrices (left to right):

```
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]  [? 0]
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[1 0]  [1 0]  [1 0]  [1 0]  [1 0]  [1 0]  [1 0]  [1 0]  [1 0]  [1 0]
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]  [1 1]
[1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 1]  [1 1]
[1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 1]  [1 1]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[? ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 ?]  [1 1]
[? ?]  [? ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[? ?]  [? ?]  [? ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
[? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [? ?]  [0 ?]  [0 ?]  [0 ?]  [0 ?]
```