Equilibrium and Termination

Vincent Danos & Nicolas Oury
School of Informatics
University of Edinburgh

We present a reduction of the Turing halting problem (in the simplified form of the Post correspondence problem) to the problem of whether a continuous-time Markov chain (CTMC) presented as a set of Kappa graph-re rewriting rules has an equilibrium. It follows that the problem of whether a computable CTMC is dissipative (ie does not have an equilibrium) is undecidable.

1 Introduction

In this note we explore an aspect of the relationship between the notion of equilibrium of a continuous time Markov chain (CTMC) and that of the traditional concept of termination in rewriting systems. Unlike in deterministic dynamical systems, a Markov chain equilibrium is not a definite state, but rather a probability over the state space which is invariant under the Markov semigroup, and satisfies an additional property explained right below.

Suppose given a CTMC with a matrix of rates $Q = (q_{ij})$ over a finite state space $I$, where $q_{ij}$ stands for the rate at which the chain jumps from $i$ to $j$. A probability distribution $p$ on $I$ is said to be an equilibrium probability for $Q$ if for all $i, j$ in $I$:

$$p(i) \cdot q_{ij} = p(j) \cdot q_{ji}$$

In plain words, this definition is saying that, at equilibrium, the probability of observing a jump between $i$ and $j$ is the same as that of observing a jump from $j$ to $i$. In some sense, time has disappeared. (The equilibrium property is called having detailed balance in chemistry, and being reversible in probability theory.) Note that for (1) to have a solution, one needs $q_{ij} = 0$ if and only if $q_{ji} = 0$. When a solution $p$ exists (and the underlying transition system is strongly connected), $p$ is the unique steady state of the chain, meaning the chain converges to $p$ no matter where it starts. The converse is not true. It is possible that the chain has a steady state which does not satisfy (1).

The existence of an equilibrium is equivalent (at least in the finite case) to the existence of an energy function for $Q$, by which we mean a real-valued function $E$ on $I$ such that for any related states $i$ and $j$,

$$\exp(-E(i)) \cdot q_{ij} = \exp(-E(j)) \cdot q_{ji}.$$ 

So, if equilibrium is the disappearance of time, exhibiting an energy function is analogous to finding a termination proof. And if one follows on the analogy, it should be possible to prove that the problem of finding such an energy function for sufficiently expressive languages of CTMCs is undecidable. This is what we do here. Specifically, we consider a class of stochastic graph rewriting systems, defined in the Kappa language, and prove that an instance $X$ of the Post correspondence problem (a simple recursively enumerable-complete problem due to Emil Post) has a solution if and only if a corresponding Kappa rule set $R_X$ is dissipative (ie admits no energy function). Essentially $R_X$ performs a general enumeration of candidate solutions for $X$. The reversibility of the search steps guarantees that no stone is left unturned.

S. B. Cooper, E. Kashefi, P. Panangaden (Eds.): Developments in Computational Models (DCM 2010)
EPTCS 26, 2010, pp. 75–84 doi:10.4204/EPTCS.26.7
The choice of the Post correspondence problem and the Kappa language makes the encoding rather simple and pleasing. (For the reader who would like to test the encoding, Kappa can be obtained at kappalanguage.org) As Kappa is used in the modelling of combinatorial biological molecular networks (as one of the languages following the rule-based approach, see eg [6–8]), this undecidability result also presents an interesting first step, and it is hoped a valuable warning sign, in the study of thermodynamically consistent rule-based models of such networks.

We start this note with a reminder of the notions of equilibrium and of the Post correspondence problem mentioned above, we then set up our encoding, and prove that it works. The conclusion will discuss further possibly intriguing consequences.

2 CTMC equilibrium

Suppose given a finite CTMC, that is to say, \( I \) a finite state space, and \( Q \) a rate matrix over \( I \) (for a complete definition see [9]).

We write \( q_{ij} \) for the rate from \( i \) to \( j \), and \( G_Q \) for the transition graph on \( I \) defined as \((i, j) \in G_Q \) if \( q_{ij} > 0 \). We suppose that \( Q \) is such that for any two states \( i \) and \( j \), if \( q_{ij} > 0 \) then \( q_{ji} > 0 \). In other words we suppose \( G_Q \) is symmetric. As said, this is a necessary condition for the existence of an equilibrium. We also write \( t(e) \), \( s(e) \) for the target and source of an edge \( e \) in \( G_Q \).

**Definition 1** The equilibrium problem for \( Q \) is to find a real-valued map \( E \) on \( I \) such that:

\[
\forall (i, j) \in G_Q : E(i) - E(j) = \ln(q_{ij}/q_{ji}) \tag{2}
\]

Such a function is called an energy function for \( Q \). It assigns, in particular, an energy difference \( \Delta E(i, j) := E(j) - E(i) = \ln(q_{ji}/q_{ij}) \) to any pair of related states (equivalently to any edge in \( G_Q \)). Depending on \( Q \), there might be no such function, or there might be many (see below).

Any map \( E \) defines a probability on \( I \) (note that we need infinite energy for \( p(i) = 0 \)):

\[
p(i) = \frac{1}{Z} \cdot e^{-E(i)} \quad \text{with} \quad Z = \sum_i e^{-E(i)}
\]

The energy/probability correspondence is a bijection between energy maps and probabilities on \( I \) - up to an additive constant for energy. Clearly, equation (2) is only a rephrasing of equation (1).

Note that: 1) if \( E \) is constant, \( p \) is uniform; 2) according to the convention chosen here (which is the usual one), if \( E(i) < E(j) \), or equivalently if \( \Delta E(i, j) > 0 \), then the equilibrium favours staying in \( i \) over staying in \( j \). That is to say the lower the energy, the more favoured the state at equilibrium.

In equation (2) we have \(|I|\) unknowns, and as many independent equations as there are pairs of edges in \( G_Q \) plus one (for normalising the distribution \( p \)) - so when do we have solutions?

**Proposition 1 (Wegscheider)** Problem (2) has a solution if and only if \( \sum_{e \in \gamma} \Delta E(e) = 0 \) over every cyclic path \( \gamma \) on \( G_Q \); this solution is unique up to a choice of one additive constant per connected component in \( G_Q \).
Suppose $E$ is a solution, then for every path $\gamma = e_1, \ldots, e_n$ on $G_Q$, one has $\sum_{e \in \gamma} \Delta E(e) = E(t(e_n)) - E(s(e_0))$. This sum is zero if $s(e_0) = t(e_n)$, that is to say as soon as $\gamma$ is a cycle (simple or not). Conversely, suppose the condition holds. For each connected component $C$ of $G_Q$ pick a node $i_C$, and an arbitrary value $E_C$; for each $i \in C$ pick a path $\gamma$ from $i_C$ to $i$ (this is possible because $G_Q$ is symmetric), and set:

$$E(i) = E_C + \sum_{e \in \gamma} \Delta E(e)$$

by the condition, this does not depend on the choice of $\gamma$, nor does it depend on the choice of $i_C$ (up to the choice of another constant $E_C$). Clearly, it is a solution. □

This condition -due to Wegscheider [12]- will be referred to as the W-condition in the sequel.

### 2.1 A simple Petri net example

Let us examine an example which will give us an opportunity to 1) extend the above definitions to a countably infinite state space; 2) introduce a simple language of CTMCs that our language of choice, Kappa, will extend later.

Consider a simple Petri net with two reversible transitions:

$$\emptyset \leftrightarrow A$$
$$A \leftrightarrow B$$

The above defines a transition system with a countably infinite state space which can be described as the set of pairs $(n, m)$ where $n$ is the number of $A$s and $m$ the number of $B$s. To obtain a CTMC we have to define rates for each of the transitions. We assume that these rates are chosen in a way that the energy differences for creating an $A$ and a $B$ are respectively $E_1$ and $E_2$ (see also the diagram below). Let us check that this CTMC satisfies the W-condition.

A cycle basis in the transition graph is formed by the following squares:

$$
\begin{array}{c}
n,m \\
\downarrow E_1 \\
n+1,m \\
\end{array} \quad \begin{array}{c}
n - 1 , m + 1 \\
E_2 \\
E_1 \end{array} \quad 
\begin{array}{c}
n - 1 , m + 1 \\
\downarrow E_1 \\
n,m + 1 \\
\end{array}
$$

where both paths have the same energy differential $E_1 + E_2$. Hence, the induced CTMC satisfies the W-condition [1]. Specifically, if we set $E(0,0) = 0$, we get $E(n,m) = (m+n)E_1 + mE_2$ which defines the limit probability:

$$p(n,m) = \frac{1}{Z} e^{-m(E_1+E_2)} e^{-nE_1}$$

There are two things worth noticing here. First, the limit probability does not depend on the rates of our pairs of transitions, but only on their ratio; this is expected. Second, the partition function $Z = \sum_{n,m} e^{-m(E_1+E_2)} e^{-nE_1}$ is bounded if and only if both $E_1 > 0$ and $E_1 + E_2 > 0$. This is new. It means that because our state space is infinite - the W-condition is not enough to guarantee the existence of an equilibrium, and one has to add the above provisions. These are rather natural as they are saying that
Equilibrium and Termination

creating As and Bs, is energetically unfavourable. If any of the conditions fail, the system creates an unbounded number of As (if $E_1 < 0$) or Bs (if $E_1 + E_2 < 0$).

Hence, thereafter, when we say that a countably infinite CTMC has an equilibrium, we mean that it satisfies the W-condition above, and, that its partition function converges.

3 PCP and the Kappa encoding

The Post correspondence problem (PCP, or PC problem) is as follows. We are given a set $X$ of pairs of non-empty words $(u_1, v_1), \ldots, (u_n, v_n)$ over some fixed alphabet $\Sigma$, and we ask if there exists $p \geq 1$ and $f : \{1, \ldots, p\} \rightarrow \{1, \ldots, n\}$, such that $u_{f(1)} \cdots u_{f(p)} = v_{f(1)} \cdots v_{f(p)}$.

As an example, consider the pairs $x_1 = (aa, a), x_2 = (ba, ab)$, and $x_3 = (b, ab)$, then:

$$x_1 x_2 x_3 = (aa, a)(ba, ab)(b, ab) = (aabab, aabab)$$

is a solution. Simple as it is, the PC problem is undecidable if $\Sigma$ has at least two symbols [11].

The next thing we do is to encode this decision problem in the W-condition of a well-chosen Kappa system. We will suppose given an instance $X = \{(u_i, v_i) ; 0 < i \leq n\}$.

3.1 Brief intro to Kappa

To this effect, we need to briefly introduce Kappa [3]. The language generalises that of Petri nets of which we have already seen an example in the previous section. One has various agent types each with a name and an associated finite set of sites. Sites can be used to bind other sites (with the restriction that any given site can be used at most once) and/or hold internal values ranging in a finite set.

Here are the four types of agents that we will need for our encoding:

- a forward agent $F(s, i)$, and
- a backward one $B(s, i)$, as well as
- a symbol agent $S(l, r, x_\sigma)$ where the site $x$ bears an internal state $\sigma$ in $\Sigma + \{\ast\}$, and
- an index agent $I(l, r, x_i)$ where $x$ bears an internal state $i$ in $\{1, \ldots, n\} + \{\ast\}$.

The objects produced by combining agents are called site graphs. One has specific rewriting rules that specify under which conditions agents bind, unbind, change internal states and get created or deleted. Rules have rates that determine uniquely a CTMC (usually countably infinite as for Petri nets) of which the states are site graphs, and of which the transitions are rule applications.

All rewrite rules will be presented graphically as this is vastly more intuitive than textual syntax where links are presented as shared exponents, internal states as subscripts, and concatenated agents as separated by a comma. E.g for a chain of symbol agents representing a word $a_1 \ldots a_n$, we would have to write the cumbersome:

$$S(l, r^1, x_{a_1}), S(l^1, r^2, x_{a_2}), \ldots, S(l^{n-1}, r, x_{a_n})$$

where the connecting sites in the chain, $l$ and $r$ have names reminiscent of ‘left’ and ‘right’. The actual integers used to identify connected sites are of no import. So instead, and equivalently, we shall use a graphical notation. Chains of symbol agents $u = a_1 a_2 \ldots a_n$ over $\Sigma$ can be represented uniquely as
indicated in Fig. 1, a notation which compares advantageously with the above one. We will elide the name of symbol and index agents, as it is easily recovered from context, and we will often do the same for site names, as they can be recovered unequivocally as well. This will permit a terse and visually pleasing presentation of the rule set encoding a PCP instance.

![Figure 1: The shorthand notation for \( u = a_1 \ldots a_n \) and its definition as an explicit chain of symbol agents. Note that the sites of the agents are understood from their position.]

3.2 Encoding

The idea of the encoding is that one starts in a state where the forward agent \( F \) holds an empty word (on site \( s \)) and a dummy index (on site \( i \)) as shown in Fig. 2. Both the dummy symbol and index are represented by \( * \) (the internal state of \( x \) is represented in the centre of each corresponding agent - purely for readability).

![Figure 2: The initial state of the \( R_X \) system; the forward agent \( F \) holds an empty symbol chain (upper dummy agent), as well as an empty index chain (lower dummy agent).]

As the computation proceeds, \( F \) concatenates to the upper symbol chain new words picked in a non-deterministic fashion in \( \{ u_1, \ldots, u_n \} \), while it records the index of these words in the lower index chain. The \( n \) corresponding rules are depicted in Fig. 3.

At some point \( F \) will switch to a \( B \) agent which will do the reverse work sliding down the index chain and re-parsing the symbol chain by chunking out words in \( \{ v_1, \ldots, v_n \} \). Importantly, the switching rule(s) as shown in Fig. 4 verifies that the index chain is not empty (ie the internal state of the lower index agent is a real index \( i \), not a dummy one). This prevents the system to switch before anything has been done.

Once it has become a \( B \), the middle agent slides backward on the index chain (which is a complete log of the choices made by the \( F \) agent), and uses it to determine which next word to recognise. This is shown in Fig. 5.

As indicated in Figs. 3–5, each of the pairs of reversible rules we have considered so far has a natural forward orientation. Indeed, the agent \( F \) natural orientation is to go \( \ldots \) forward and extend the chains or eventually switch to the \( B \) form, while the agent \( B \) natural one is to consume the symbol chain and
Figure 3: \( F \)-rules in \( R_X \) (with forward orientation from left to right): the forward agent \( F \) extends the symbol chain by picking a word \( u_i \) and records the index \( i \) in the index chain. Note that in the rule left hand side the current rightmost symbol and index agents need not be fully detailed (this basic Kappa convention is sometimes referred to as the ‘don’t care, don’t write’ principle [6]). Note also that the shape \( u_i \) is a (non-empty) chain, not a single agent (as defined in Fig. 4).

Figure 4: Switching rules in \( R_X \) (with forward orientation from left to right): under the condition that the \( F \) agent has a non empty chain (the index one or equivalently the word one) it can flip into the \( B \) agent.

Figure 5: \( B \)-rules in \( R_X \) (with forward orientation from right to left): the backward agent \( B \) re-parses the symbol chain backward by picking a word \( v_i \) as indicated by the index \( i \) it is bound to in the index chain.

Figure 3:

Figure 4:

Figure 5:

Go backward. We refer henceforth to this natural direction as the forward direction. By construction, going backward (aka backtracking) is deterministic, and any trace starting from the initial state can be visualised as an exploration of the PCP exploration tree, where backward steps allow backtracking and therefore guarantee at all times a path to any solution if there is one (possibly with infinite average hitting time). (This systematically available backtracking is reminiscent of the reversible CCS formalism...
of Ref. [2]). In particular, agent $B$ can backtrack only by recreating the chain $v_i$ it has erased. It is the index chain/log that forces this. As a consequence, according to the rule set $R_X$ defined so far, $B$ cannot switch back to $F$ in any other state than the one at which $F$ itself switched.

From this, it is easy to see that success, meaning a backward agent with an empty symbol chain, is equivalent to finding a solution to PCP.

In fact, supposing all the rules in $R_X$ have non-zero rates we get:

**Proposition 2** The set of solutions of a PCP instance $X$ is in bijection with the successful configurations which the rule set $R_X$ can reach from the initial state.

As $B$ does not erase the index chain (see Fig. 5), a successful configuration (by definition one where agent $B$ has an empty symbol chain) contains a lower index chain which is a solution of the $X$ instance.

### 3.3 Undecidability of the W-condition

Now that we have a neat embedding of the search for solutions to $X$ as a rule set $R_X$, we need to relate its success to the equilibrium problem. The idea is the following. Because of the earlier remark on the deterministic nature of reverse steps, any rewrite trace is equivalent to a purely forward one, up to trivial cancellations. Thus, we are at liberty to add a series of new rules for $B$ to consume also gradually the index chain once a success has been recorded.

These new rules are shown Fig. 6–7. We call $R'_X$ the rule set formed by $R_X$ together with the new rules, and assign (non-zero) rates to all rules in $R'_X$ so as to obtain energy differences that are zero, except for the second switching rules of Fig. 7, where the energy difference is set to a constant $E \neq 0$.

![Figure 6: Index chain deletions in $R'_X$ (with forward orientation from left to right): the $B$ agent, once a successful configuration is reached (as can be seen from the fact that the upper symbol chain is empty), progressively erases the index chain in order to return to the initial state.](image)

If $X$ has a solution, then one can simulate its discovery by a purely forward trace, which one can then conclude using the additional rules to return (in a forward way) to the initial state. This means there is a forward cycle in the state space. By construction its energy is $E \neq 0$, which is a violation of the W-condition for $R'_X$, within the connected component of the initial state $C$ (thereafter called simply the initial component).

Conversely, suppose one has a violating cycle in $C$. The earlier remark about rewriting traces up to trivial forward/backward cancellations still applies with our bigger rule set $R'_X$. By definition, such cancellations
Equilibrium and Termination

Figure 7: Second switching rules in $R'_X$ (with forward orientation from left to right): the B agent, once the index chain is erased (but for one remaining agent), flips into the F agent, returning this to the initial state, and creating a loop conditioned on the existence of a success. The coefficient $E$ indicates the energy difference of the second switching rules (for the left-to-right direction).

do not change the energy difference associated to the path, nor the fact that it is a cycle. So we may assume our violating cycle $\gamma$ has no such cancellations. For $\gamma$ to violate the W-condition it must go through one of the second switching rules, as these are the only rules with a non zero energy differential. This means that one can take the origin of $\gamma$ and choose its orientation in such a way that $\gamma$ starts forward from the initial state and finishes with a second switching rule. But then $\gamma$ must attain a successful configuration.

We have proved:

**Proposition 3** A PCP instance $X$ has a solution if and only if the rule set $R'_X$ violates the W-condition.

3.4 Undecidability of dissipativity

This proposition is not yet as strong as one would like. As we have seen earlier, for countably infinite state spaces, the W-condition is not enough to ensure an equilibrium, as the associated partition function might diverge. So we have not obtained yet the stronger result that $R'_X$ is dissipative (ie does not have an equilibrium) if and only if $X$ has a solution.

Worse, with the particular chosen rates, this is clearly wrong. Let us see why. Reconsider the case where $X$ has no solutions. In this case the connected component $C$ of the initial state does not contain any success state, and therefore, from the point of view of $C$, the CTMC is entirely described by the rule set $R_X$. Since we have assigned a zero energy difference to all rules in $R_X$, every state in $C$ should have the same probability (conditioned on the initial state being in $C$), and since $C$ is countably infinite, this is absurd and therefore can only mean that $Z$ diverges.

So, to get a convergent $Z$ we need to tweak our assignment of energies. It turns out that there is a very natural way to do this which is continuous with our previous construction. Pick a real number $\varepsilon$, this will be our quantum of energy. Assign to any state $x$ in $C$ the energy $n \cdot \varepsilon$ where $n + 1$ is the length of the index chain of state $x$ (equivalently $n$ is the number of non dummy indices in $x$’s index chain).

Write $R'_X(\varepsilon)$ for this more general assignment. The former energy assignment corresponds to $\varepsilon = 0$, ie $R'_X = R'_X(0)$. Except for the second switching rules, all transitions can be made compatible with this assignment (as stipulated in equation (2)), as each induces a variation of the length of the index chain.
which is well defined. (For the second switching rules to be compatible, one would need \( E = -\varepsilon \).) So, clearly any cycle that does not use a second switching rule still has zero energy differential.

As in the special case \( \varepsilon = 0 \), we reason that if there is a solution to \( X \), there must be a \( W \)-violation, hence no equilibrium. Now, if there is no solution to \( X \), the second switching rules are never used, else the \( W \)-condition is satisfied (by the point made just above). At this stage, we have recovered the \( R_X'(0) \) argument. But this time we have more.

**Lemma 1** Define \( \Omega_n \) as the set of states in the initial component \( C \) with energy \( n \cdot \varepsilon \). If \( X \) has no solution, then \( |\Omega_n| \leq (n + 1)|X|^n \).

If \( X \) has no solution, any state in \( C \) with a given index chain of length \( n + 1 \) is either a unique state with \( F \) as the middle agent, or one of at most \( n \) states with \( B \) as the middle agent, depending on how far \( B \) has slid back on the index chain. As there are \( |X|^n \) such chains, the upper bound follows. □

Hence, if \( \varepsilon > \log |X| \), the associated partition function over \( C \) converges:

\[
\sum_{x \in \Omega_n} e^{-n\varepsilon} \leq (n + 1)|X|^n e^{-n\varepsilon} \sim ne^{-n(\varepsilon - \log |X|)}
\]

This implies:

**Proposition 4** Let \( X \) be a PCP instance, and suppose \( \varepsilon > \log |X| \), then \( X \) has a solution if and only if the rule set \( R_X'(\varepsilon) \) is dissipative.

Hence the problem of whether a countably infinite computable CTMC is dissipative is undecidable.

In passing, \( \log |\Omega_n| \approx n \log |X| \) is referred to in statistical physics as the *entropy* of the energy equivalence class \( \Omega_n \) (note that the entropy is a macroscopic notion that presupposes a macro-observable - here the energy). So one can view our argument as saying that by fixing \( \varepsilon > 0 \) sufficiently positive, the entropy term can be controlled by the energy one. This is what physicists call a phase transition. In effect, all we need is to set a sufficient energy penalty on the exploratory behaviour of \( F \) - that is to say its forward moves- to make the probing of longer potential solutions increasingly more expensive, and therefore more unlikely. In this argument, we have chosen a uniform penalty \( \varepsilon \) per increment of the index chain, but one could let \( \varepsilon \) depend on the chain length. (This leads to a probabilistic version of König’s lemma, where the branching degree of the forward agent can be countered by a decreasing likelihood of exploring a branch.)

## Conclusion

We have proved that the problem of whether a countably infinite computable CTMC is dissipative is undecidable. Early on (in §2.1), we have described an example using a simple Petri net. Despite being complex, reachability is decidable for Petri nets and so they cannot host an encoding of PCP similar to the one we have used here. Nevertheless, there should be a refined version of the result that we have presented that would explain how difficult it is to determine whether a given Petri net is dissipative. One might speculate that this latter problem is at least as difficult as reachability. Likewise, it would be interesting to derive an NP version of our result. As bounded PCP (where the length of the solution is bounded at the outset) is NP-complete, one might think of using the same basic setting. This prompts another question. As said in the introduction, the PCP/Kappa couple does not play a fundamental role
Equilibrium and Termination

here. It is a way to get a precise formulation of the problem. It could be instructive to attempt to repeat this argument at a more abstract level by using an axiomatic treatment of stochastic rewrite systems.

Another more practical research thread that is suggested here is directly related to the modelling issues at the heart of Kappa [6] and similar rule-based languages with a CTMC semantics such as the BNG one [1, 4]. It is the question of finding tractable forms of the W-condition that would be sufficient to ensure equilibrium (but obviously by our very result, not necessary). In the context of Kappa it is natural to think of introducing a class of energy functionals that would guarantee stronger and hopefully more feasible forms of the W-condition - perhaps based on the usage of local patterns as is customary in Ising models and derivatives thereof (eg see [10, Chap. 12]). This is a problem of static analysis that we intend to investigate in the near future. Whichever structure one chooses to achieve this, it seems that in the context of model fitting, which is of cardinal importance in biological modelling (eg see Ref. [5]), our result establishes that thermodynamic consistency has to be “wired in” the framework and can hardly be an afterthought.

References


