Greg Kuperberg's Lectures on

Introduction to Quantum Computation

Sanchayan Dutta
dutta@ucdavis.edu

Contents

1 Lecture 1 (12 November 2021)

1.1 Fundamental Goal

To characterize the category of realistic maps between states.

We have two conclusions:

1. If a map $E\colon A^{\#} \to B^{\#}$ is realistic then it is TPCP. If E is not linear it violates classical superposition. If it's not CP then it and together with a companion thing can create negative probabilities or non-real probabilities. If it is not TP is doesn't preserve probability.

2. (**Stinespring-type theorem**) If E is TPCP, then you can produce it as a composition of realistic components. In fact not very many of them – just a factorization consisting of ancilla pure states on a Hilbert, followed by the transpose of an algebra homomorphism.

Say that $A = M(a)$ and $B = M(b)$ then (2) is equivalent to a structure theorem due to Krauss for the structure of E. At a conceptual level this structure theorem plays an important role and in this way of doing things quantum probability (and quantum superposition) arises as a corollary of classical probability (and classical superposition).

Kraus' theorem (named after Karl Krauss) characterizes CP maps that model quantum operations between quantum states. Informally, the theorem ensures that the action of any such quantum operation E on a state ρ can always be written as $E(\rho) = \sum x\rho x^*$ for some set of operators $\{x_k\}_k$ satisfying $\sum_k x_k^* x_k = \mathbf{1}$ where $\mathbf{1}$ is the identity operator.

$E(\rho) = \sum x\rho x^*$ is a classical superposition of terms. If ρ is a pure state of the form $|\psi\rangle\langle\psi|$ then $x|\psi\rangle\langle\psi\rangle x^*$ is a linear operator on the pure state. The $x\rho x^*$ part is a quantum superposition.

Note: Here in $E\colon A^{\#} \to B^{\#}$, A and B need not necessarily be qudits. They can be semi-quantum in various ways.

In this way, we get our category QProb and the finite-dimensional part qProb.

1.2 QProb and qProb

Theorem: If E (in QProb) is *reversible*, E^{-1} exists and is also TPCP, and E comes from an algebra isomorphism. This means there's a restricted way to go backwards. The only way this can happen is if all of the algebra apparatus is preserved for Alice and Bob.

(In terms of category theory, if you have any category with a set of reversible maps, the maps are said to have inverses in that category.)

Theorem: Any algebra isomorphism between qudits or $\mathcal{L}(H_A)$ and $\mathcal{L}(H_B)$ is given by a unitary operator. This is another example of getting quantum superposition and quantum linearity out of classical superposition and classical linearity.

(In pure math, if C is a category and $f\colon A \to B$ in the category C has an inverse f^{-1} also in the same category, then f is called invertible. Well, f is also called an isomorphism. In physics one says, f is reversible. E may represent the evolution of system. We may ask when the time reversal is well founded under the laws of physics.)

An example of TPCPs: Unitary operators and isomorphisms and more generally, automorphisms.

Every von Neumann algebra has unitary elements. In any VNA M, there is a unitary subgroup $U(M)$ consisting of the solutions to $u^*u = uu^* = 1$. If $M = \mathcal{L}(H)$ then the corresponding algebra automorphism is $x \to uxu^*$. Physicists and QIT folks will call this transformation unitary whereas operator algebraists will call this kind of automorphism inner.

- $U(M) = M(S^1)$ which are the circle-valued observables
- $U(M)$ is friends with $M_{\mathbb{R}}$, $M_{\mathbb{Z}/2\mathbb{Z}}$, etc.

- $M_{\text{normal}} = \{z \text{ s.t. } zz^* = z^*z\}$ = one-shot measurements. The order in which you measure the real and imaginary parts of z does not matter since they commute with each other. See: physics.stackexchange.com/a/82616 for more details. Unitary is a special case of normal.
- $U(M)$ represents measurements that take values in a unit circle on the complex plane.
- If M is commutative, the only automorphisms $x \to uxu*$ is the identity.

1.3 Modelling Measurements using TPCPs

(1) **Automorphisms and isomorphisms**

(2) **Visible measurement**

M = any system A = a classical observer, for starters a bit

Given a homomorphism f from A to M there is then a measurement TPCP.

$E\colon M^{\#} \to (M \otimes A)^{\#}$ where afterwards M holds the posterior state and A holds the measured value. We can suppose that A injects into M. So we need a algebra homomorphism and we can suppose that it is injective. If it's not injective then the measurment will have redundancies. In some way or another you will be measuring the same thing multiple times.

If A is a finite digit with a configurations, you can take E apart into a disjoint booleans $b_1, \ldots, b_a$. Disjoint means $b_j b_k = 0$ for $j \neq k$. So they are simultaneously measurable. Then the posterior state in M is a sum over the posterior state formula from before:

$$E(\rho) = \sum b_a \rho b_a \otimes [a]$$

where $[a]$ is the output value. This is the qudit case for M and $b_k = f(k)$.

(3) **Discard or Departure**

We have the final system $\mathbb{C} = M(1)$ and for any other A, there is a unique TPCP $A^{\#} \to \mathbb{C}^{\#}$ interpreted as discarding A or A leaves the room.

Final and Initial Objects

If C is a category, a final object is one that has exactly one morphism from any other object. An initial object has exactly one morphism to any other object. All final objects are isomorphic. Also all initial objects. Suppose A and B are two final objects. Then there can be only one map from A to B and only map from B to A.

Example: The category Set has both a final object and an initial object and they aren't isomorphic to each other.

Think of a set such that there is always only one function from it. The empty set is initial. Think of a set such there is only one function to it. 1-point sets are final.

Example: The category Vect. Think of a vector space such that there is only one linear map from it. The 0-dimensional vector space $\{0\}$. Now think of a vector space such that there is only one linear map to it. The 0-dimensional vector space again! So this vector space is both initial and final.

So QProb has a final object. The VNA $M = \mathbb{C}$ representing the complex numbers. QProb doesn't have an initial object. Classical probability too doesn't have an initial object but it does have a final object. Which means there is only one way to "destroy" but many ways to "create". This is because (say) you can create a qudit in any case, but there is only one way to destroy/discard a qudit.

In fact, the maps from $\mathbb{C}^{\#} \to A^{\#}$ are a copy of A^{Δ}. You can reproduce the states on A as TPCPs from non-existence to A.

Measurement Process

Combining (2) and (3), measurement into A followed by discarding A is the hidden measurement TPCP. So Alice can come in, measure (using a homorphism) and leave. She will never tell you what was measurement. If the system M was non-commutative, it's state would have changed.

If M is a qudit and A is finite, $\rho \to \sum b_k \rho b_k$.

E.g. If M is a qubit and the measurement is boolean (2-valued) then the hidden measurement E collapses the Bloch ball to an axis.

We've described **perfect measurements**, given by homomorphisms from A to M. If f is not injective, then we can replace A by $B = f(A)$. In this context, first Bob measures M and then Alice measures Bob. The only way this is possible is if A has unused values. The actual measurement is from 1 to 6. But Alice register holds values from 1 to 10. So you can simplify the measurement so that the register only holds values from 1 to 6.

If $M = M(d)$ is a qudit then a classical $A = a\mathbb{C}$ only embeds in M when $a \leq d$.

Another description: If a qudit $M(d)$ or a larger $L(H)$ has a Hilbert space H and A is the finite register of a perfect measurement then the measurement comes from an orthogonal decomposition of H into subspaces H_k (k in A).

Say, $7 = 3 + 3 + 2$ (ternary measurement). 7-dimensional Hilbert space. Choose any 3-dimensional subspace. Projection onto that is the first boolean. In the 5-dimensional complement choose another 3-dimensional subspace. The remaining complement is 2-dimensional.

(Note that each boolean element of the algebra have a rank which corresponds to the dimension of the image space.)

Here b_k = perpendicular projection onto H_k. $|\psi\rangle \to b_k|\psi\rangle$ is also the unnormalized Bayesian posterior of b_k.

Imperfect measurements aren't given by homomorhisms at all. In the qudit case they are called POVMs. POVMs are measurements that allow noise blur.

We note that A can be *any set*. It doesn't have to a subset of $\mathbb{R}$. Real-valued measurements are used to help answer what is the system; not what measurements can we do. Stereotypical physicist says that they can't tell you whether it's an apple, banana or pear unless your number them! That doesn't make sense in the context of measurements.

If x is a real spectrum discrete operator then we can let A be the spectrum and then get a number-valued measurement where H_k is an eigenspace.

A TPCP from $A^\#$ to $B^\#$ might also be unital if we use trace to identity A with $A^\#$, say in the qudit case. If A and B are qudits and E is a TPCP $A^\# \to B^\#$ then the unital condition $E(1) = 1$ is equivalent to E preserving the uniform state, the center of A^Δ.

Classically E (regardless) is a stochastic matrix and then this extra makes E doubly stochastic. Quantumly, such an E is still called doubly stochastic.

What does a doubly stochastic map do? It doesn't make any negative probability, preserves total probability and most importantly preserves the uniform distribution. To preserve uniform distribution the row sums should also be 1 (in addition to column sums being equal). E.g. E can be a permutation matrix which is exactly doubly stochastic and deterministic.

1.4 Birkhoff's Theorem

Theorem (Birkhoff): Every doubly stochastic matrix is a convex sum of permutation matrices.

Example: $A = B =$ a bit, $E =$ bit flip with probability p

$$E = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

If $A = B = M(d) =$ a qudit then unitary operators are doubly stochastic.

(1) Any convex sum of them is doubly stochastic and many TPCPs aren't. Visualize this in qubit case. An unitary or algebra automorphism will simply rotate the Bloch ball. The middle will still go to the middle. But then you can also throw away the qudit and recreate a state at the pole. It's perfectly realistic but not a convex sum of unitaries.

(2) The **quantum Birkhoff theorem** is true for qubits, any UTPCP (doubly stochastic) is a convex sum of unitaries but it's not true for qudits with d at least 3. The real flavour of Birkhoff's theorem is that it's a complete description of the extremals.

The question of extremal points of TPCPs is a notorious problem. Unitaries are simple examples of extremal UTPCPs. The extremal CP rays are easy – they are individual Kraus terms.

(3) Hidden measurements are convex sums of unitaries.

Decoherence due to a hidden measurement can be accounted for by multiplying each subspace with a scrambled phase factor.

1.5 Bell's Theorem

Theorem (Bell): QProb does not embed in Prob as a tensor category, not even approximately. Whereas Prob is immediately a subcategory of QProb. Anything like an embedding from QProb to Prob would respect the Bell inequality for two qubits.

Any category can be realized as a category of sets and functions between them. But that theorem is false if you state tensor categories. To put QProb in Prob you would have to repeal the concept of joint system and have a grand theory of codependence. Then you might be able to wish away quantum probability. Quantum entanglement even for two qubits cannot be modelled by classical systems, even infinite ones. You would have to able to split up a system into Alice and Bob to even have a theorem like this.

If you have just one quantum computer that you don't split into pieces you *can* model that using classical system. Though if you do the corresponding classical might be exponentially larger. By the way, we should note that closed system rules don't apply to measurement.

Note: The semi-quantum trit has more than one possible center, depending on what you're trying to do. The centroid may not be the maximum entropy point.

1.6 Models of Computation

Any tensor category supports circuits = Tensor networks with a time arrow.

Three fundamental cases for us:

1. **Deterministic circuits**: The category is Set. The tensor operation is Cartesian product. (AND, OR, and NOT are present here.)
2. **Randomized circuits**: The category is Prob. The tensor operation is tensor product of commutative algebras.
3. **Quantum circuits**: The category is QProb. The tensor operation is tensor product of the finite dimensional VNAs.

In (1), certainly bits $\mathbb{Z}/2\mathbb{Z}$ or 2-element sets in general are objects. AND, OR, NOT, and COPY (1 bit goes in and 2 bits go out) are morphisms. The first theorem as a tensor category, is that these gates generate a big part of Set (not all objects but many objects, and all morphisms). In fact, they generate all objects with 2^n elements and all morphisms in between.

Extra trick, the Karoubi envelope trick for any category. Any category C has a Karubi envelope, where for each object A and each idempotent $f^2 = f$ on A, (A, f) becomes a new category.

If we throw in the Karoubi envelope as part of "generate", the standard gates (*) generate all of Set.

The purpose of this is to make sets of each finite size from just sets whose size is a power of 2. At first, I can only make sets whose sizes are powers of 2. How do I make a set with 3 elements? Well, choose a mapping from 4-elements (2 bits) to 4-elements (2 bits) such that that the stray 4th value to send it to one of the other three. So you empower this 4-element set with a correction demon and the correction demon just sends the 4 bitstring values back to 0.

$$\text{Karoubi Demon}: 00, 01, 10, 11 \to 00, 01, 10, 00$$

Register coerces the values to trit values. Demon sends $3(= 11)$ back to $0(= 00)$.

It's a way to make new objects from old objects in any category. This is a object factory.

Same principle as for finitely generated group. That the finite generating set of the category set doesn't matter much. You can interconvert with a finite overhead.

P/poly = non-uniform polynomial time is a set of functions (or sequences of them) that have poly-sized circuits.

For randomized computation, there's a problem that's going to continue in the quantum case. The set of morphisms $\mathrm{Hom}(A, B)$ is uncountable but a finite set of gates can at best reach, at best, a countable number of them. The morphisms are stochastic maps and have continuous parameters.

Two solutions:

(1) (Less popular) Allow a continuous family of gates.

(2) (More popular) Allow circuits to approximate rather than equal a target stochastic map, i.e., densely generate.

Produce for me a coin flip such that the probability of heads is $\frac{1}{e}$. With fair coin flips you can only create diadic rationals where denominator is some power of 2.

Theorem (AND, OR, NOT, COPY and "RANDOM" = random bit creation, 0-ary gate with 0-input and 1-ouput), together densely generate Prob. We still Karoubi envelope trick as we to generate objects whose dimensions are not powers of 2.

This lets you slide back from Bayesianism to frequentism. The random bits can be viewed as certificate that are helping the computation (frequentist interpretation).

2 Lecture 2 (19 November 2021)

2.1 Measures of Fidelity

$E\colon \mathcal{A}^{\#} \to \mathcal{B}^{\#}$ is a desired quantum map. You might instead see $F\colon \mathcal{A}^{\#} \to \mathcal{B}^{\#}$.

1st for states $\rho, \sigma \in \mathcal{M}^{\triangle}$:

$$d(\rho,\sigma)\colon = \max_{b\in\mathcal{M}_{\mathbb{Z}/2\mathbb{Z}}} [\rho(b)-\sigma(b)] \overset{\text{Thm}}{=} \frac{1}{2}||\rho-\sigma||_1$$

This is called trace distance, infidelity (sort of) or variation distance.

$\frac{1}{2}||\rho-\sigma||_1 \overset{\text{Thm}}{\Longrightarrow}$ Same bands for general distinguishability for ρ vs. σ or E vs. F for any use with only one copy.

$d(E,F)\colon = \sup_{\rho\in\mathcal{A}^{\Delta}} d(E(\rho),F(\rho)) \overset{\text{Thm}}{\Longrightarrow}$ Same bands for general distinguishability for ρ vs. σ or E vs. F for any use with only one copy.

Worst case infidelity,

Theorem $d(E\otimes G, F\otimes G) = d(E,F)$. Contrast TPP vs. TPCP. Also contrast ensemble fidelity vs. entanglement.

$$d(E_1 \otimes E_2) \leq d(E_1,F_1) + d(E_2,F_2)$$

$$d(E_2 \circ E_1, F_2 \circ F_1) \leq d(E_1,F_1) + d(E_2,F_2)$$

2.2 Karp-Lipton Theorem

Karp-Lipton Theorem: $\mathsf{P/poly} = \mathsf{P}_{\text{non-uniform}}$

$\mathsf{P/poly}$ represents a Turing machine with a polynomial time budget and and polynomial advice from an angel. $\mathsf{P}_{\text{non-uniform}}$ represents sequences of poly-sized circuits.

The $\supseteq$ containment is thought of as angel providing the circuit whereas the $\subseteq$ containment is an unrolling argument.

Theorem: $\mathsf{P} = \mathsf{P}_{\text{uniform}}$

P represents polynomial sized Turing machines whereas $\mathsf{P}_{\text{uniform}}$ represents circuits drawn by one polynomial-time algorithm.

The $\supseteq$ containment is thought of as simulating your own circuit. The $\subseteq$ containment is an unrolling argument.

2.3 Tensor Circuits

Tensor networks in suitable $\otimes$ category gives you circuit computation:

Objects	Maps	$\otimes$	poly-sized circuits
Set	functions	$\times$	P/poly
Prob	stochastic	$\otimes$	BPP/poly
QProb	TPCP	$\otimes$	BQP/poly

Fact: In all 3 cases, you get correct P, BPP or BQP in one of two ways:

1. TM draws a circuit. 2. Use periodic circuits (special case of 2) = cellular automata (Fig. 1)

Each category has generating sets, except, for Prob and QProb you need dense generation. You need a Karoubi construction (make new objects with a Karoubi coercion idempotent map) to get all objects instead of just $(\mathbb{Z}/2\mathbb{Z})^n$.

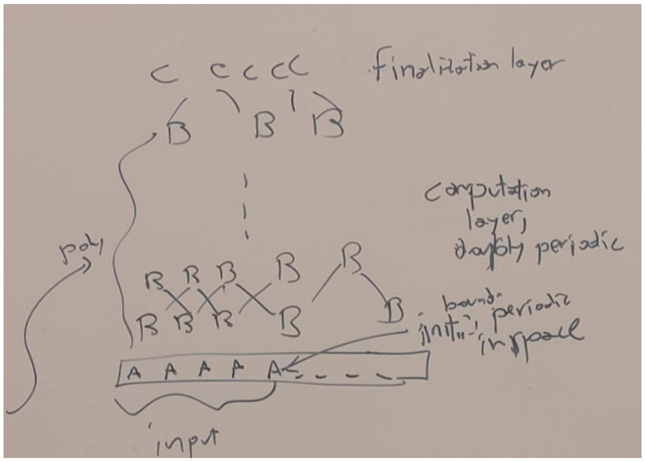

Figure 1: Periodic Circuit

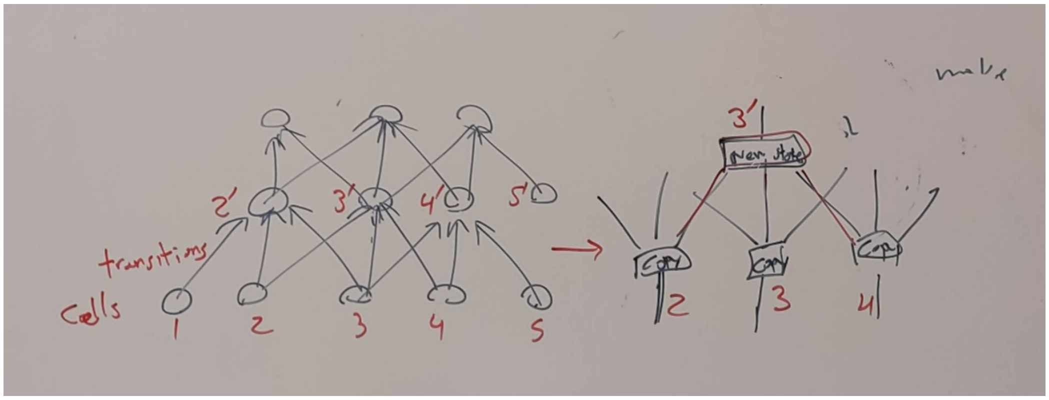

Figure 2: Computation by an Automaton

P/poly: $\mathbb{Z}/2\mathbb{Z}$, AND, NOT, OR, COPY is used to generate all objects but it doesn't matter.

BPP/poly: Random source factorization. We determine gates to any 0-ary random bit gate = generator. (**Theorem**: This kind of generation is dense.)

BQP/poly: Stinespring dilation. Promote all bits to qubits (except at the end!). Promote all TPCPs to unitaries + initialize fresh ancilla qubits in $|0\rangle$ states.

We said dense generation. In both cases, there is the *efficient* dense generation problem. To express my gates in your gates you need larger and larger approximate circuits and that should be uniform.

Necessary condition: The parameters in the gates should be efficiently computable numbers. Then there's a **theorem** saying that efficient generation is possible. And, there exists infidelity with a polylog(ϵ) overhead. This is basically the statement of the **Solovay-Kitaev theorem** in the quantum case.

2.4 Computation by Automaton

Figure 2 illustrate of how computation is done by an automaton.

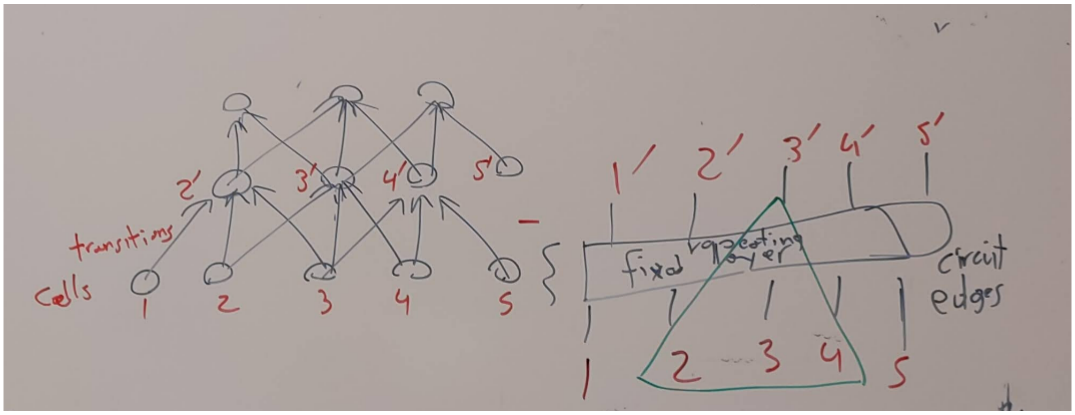

Figure 3: Representing Periodic Circuit as an Automaton

3 Lecture 3 (30 November 2021)

3.1 BQP (Bounded-error Quantum Polynomial)

1. Functions **computable by polynomial sized circuits** (uniform families or even just periodic circuits) of TPCPs.
2. **Circuit cleanup** - Using gate-by-gate Stinespring dilations can reduce to
a) Unitary gates
b) Ancilla initialization
c) Measurement at the end
3. Reasonably intuitive, realistic extension of (2) is a **classical TM with a quantum tape**.

The B, by analogy with BPP means bounded error probabilistic. Shoehorns Prob/QProb models into framework of deterministic questions. Input is classical deterministic.

$$\Pr(\text{Correct answer at the end}) > \frac{2}{3}\ (\text{say})$$

It's even better to say $> 1-\epsilon$. $\text{poly}(|\text{input}|, \log(\epsilon))$.

3.2 QFT vs. DFT

QFT vs. DFT on $\mathbb{Z}/2^n\mathbb{Z}$ Input: n qubits vs. 2^n floats. Performance: $\text{poly}(n)$ vs. $\mathcal{O}(N \log N)$ where $N = 2^n$.

$$\mathbb{Z}/2^n\mathbb{Z} \hookleftarrow \mathbb{Z}/2^{n-1}\mathbb{Z} \hookleftarrow \mathbb{Z}/2^{n-2}\mathbb{Z} \hookleftarrow \mathbb{Z}/2^{n-2}\mathbb{Z} \hookleftarrow \dots$$

3.3 Basic Complexity Classes

Figure 4 shows a chart of some fundamental classical and quantum computational complexity classes.

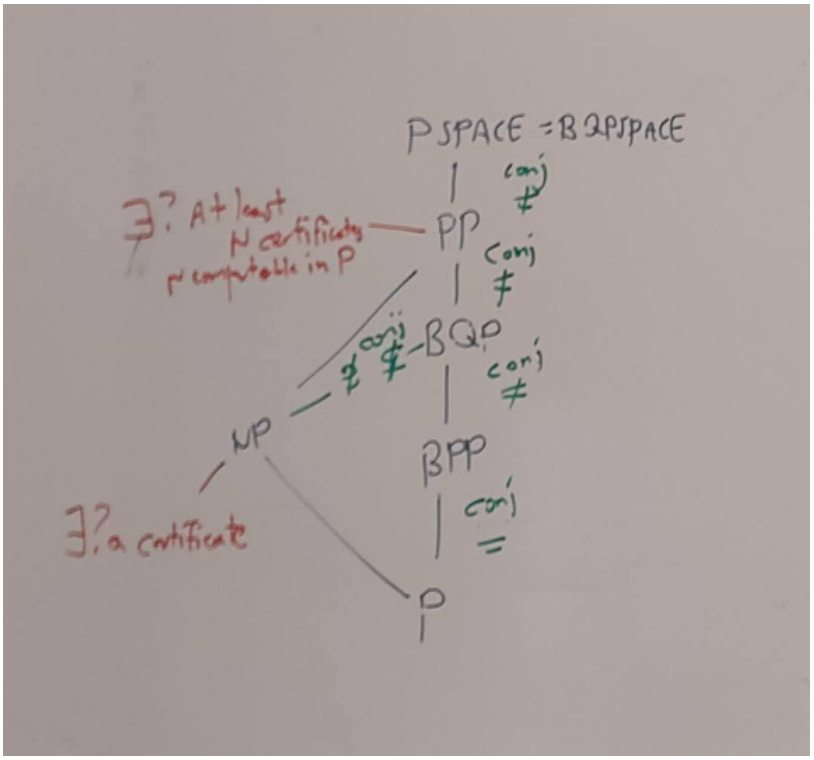

Figure 4: Some Fundamental Computational Complexity Classes